Population structuring in southern African zebras

Yoshan Moodley

A thesis presented for the degree of Doctor of Philosophy at the Wildlife Genetics Unit, Department of Clinical Laboratory Sciences, Faculty of Health Sciences, University of Cape Town.

Cape Town, September 2002

Declaration

I declare that this thesis is my own, unaided work. It is being submitted for the degree of Doctor of Philosophy in the University of Cape Town. It has not been submitted prior to this for any degree or examination in any other University.

Signed by candidate

this 26th day of SEPTEMBER 2002

for Deiffenbaker, Winnepeg and Fatboy

Abstract

Zebras, the most diverse and successful of the equids, have undergone recent and in some cases drastic population reductions as a result of the actions of man. This situation is epitomised in southern Africa, the most infrastructurally-developed sub-region of the continent, where two of the three extant zebra species occur. In this study, the genetic information at 15 microsatellite loci was obtained from 295 mountain zebra (*Equus zebra*) and 184 plains zebra (*Equus quagga*) samples, all from a total of 24 populations in order to ascertain the factors shaping intraspecific population structure and to identify areas of immediate conservation relevance. The control region of the mitochondrial genome was sequenced in a sub-sample of each population in order to determine the haplotypic structure of populations, phylogeography and to estimate the coalescence of intraspecific lineages. The genetic heterogeneity in both specific data sets was significantly correlated with the geographic distance between populations, implying that isolation by distance was the basis of population structure in southern African zebras. Among mountain zebras, the highly vulnerable Cape mountain zebra (*E. z. zebra*) could not be distinguished from its Namibian sister subspecies, the Hartmann's mountain zebra (*E. z. hartmannae*) at either molecular level. The three surviving aboriginal stock populations of Cape mountain zebra are all highly inbred as a result of extreme population bottlenecking and accelerated genetic drift. However, in many cases, these populations have become fixed for different alleles, which artificially biased measures of differentiation upwards. The conservation implication of this is that the entire Cape mountain zebra metapopulation still contains much of its historical levels of genetic diversity, which if intensively managed, may be maintained in newly-seeded heterozygous populations. The free ranging and abundant Hartmann's mountain zebra was relatively unaffected in the central and northern parts of its Namibian range, although the data presented show that this situation may change if some of the recommended management options are not implemented. In southern African plains zebra, subspecific population structuring between the Damara zebra (*E. q. antiquorum*) and Chapman's zebra (*E. q. chapmani*) was not supported. *Equus quagga* was more genetically diverse than *E. zebra*, with almost all wild plains zebra populations being characterised by high genetic variation. Population reductions at the turn of the century resulted in a significant loss of diversity in KwaZulu plains zebra populations. It is also recommended here that the many small populations of plains zebra in South Africa be managed to maintain adequate levels of genetic diversity.

Acknowledgements

I thank my supervisor, Eric Harley for his advice, insight and unfathomable wisdom. Also of particularly special mention is Ingrid Baumgarten, to whom I am eternally grateful for her part in getting the project off the ground and for numerous tips and pointers in the laboratory. For help in the tedious task of manually scoring microsatellite autoradiographs I thank, apart from the above-mentioned, Cheryl Johnson, Pete Berman and Yumna Dhansay. Also of great help in the laboratory were Michael Cunningham, Abdu Mohamed, John Thornthwaite, Carlos Ruiz and David Hasberger.

Others who have helped with advice and knowledge were Reinhold Rau of the South African Museum, Cape Town and Peter Lloyd of the Scientific Services Division of the Western Cape Nature Conservation Board. In this respect Banie Penzhorn and Peter Novellie are also thanked.

In the collection of Cape mountain zebra samples I wish to thank Pete Morkel and Paul Meyer formerly of the South African National Parks Game Capture Unit; expert SANP helicopter pilot Martin Louw; Gail Cleaver, manager of Kammanassie Nature Reserve; Tom and Leisl Barry, managers of Gamka Mountain Nature Reserve; Gundilein Miethe for fearlessly and tirelessly chasing zebras; Nicky and Chris Hess, Melissa and Ben Swanepoel of DeHoop Nature Reserve; game capture specialist Hennie Barnard; Peter Burdett of Karoo Nature Reserve; Pierre Nel of Free State Nature Conservation; Etienne Fourie of Bontebok National Park, game capture vet Hamish Currie; Johan de Klerk of Mountain Zebra National Park; Guy Castley and Mike Knight for SANP collection permits; Fritz Rohr for making biopsy darts and Martin Haupt for the loan of the Telinject dartgun; Mr and Mrs van der Westhuizen of the farm Klipdrift and Paul Heyns of the farm Mount Hope are thanked for pictures of the extinct Outeniqua mountain zebras.

While collecting Hartmann's mountain zebra samples I have indebted myself to Louis Geldenhuis for advice, Kevin Davidow of the Nakara Tannery, Gabbie and Mr Nieberle of Western Trading Corporation, Mr Gorn of Trophaiendienste, the proprietors of

Otjiwarongo Taxidermy, Piet Burgers at the MET office in Otjiwarongo and Nigel Berryman at the MET office in Okaukuejo. Pauline Lindeque and Holger Kohlberg of the Specialist Scientific Division of the MET were extremely helpful in the processing of essential collection permits.

For help in the collection of plains zebra samples I thank Winnie da Costa of Trans African Taxidermists, Randburg; Graham Thompson; Tracy Swann of Bromley Game Skins, Harare; Debbie Peake of Mochaba Safaris, Maun; Deon Hignet for processing import permits and Peter Taylor and Karen Bishop for collecting all the KwaZulu samples.

Much appreciated help with the manuscript was received from Jack Harington, Lynda Murray and Neville Pillay.

For financial support I wish to thank the National Research Foundation of South Africa and the Deutscher Akademischer Austausch Dienst.

Lastly I would like to thank my parents and all others who supported me personally.

List of Figures

List of Tables

Abbreviations and acronyms

A	Allelic diversity, the mean number of alleles per locus
A, G, C, T	Nucleotide residues in DNA sequences
AGaR$_{ST}$	program for the Analysis of G_{ST} and R_{ST}, Harley (2002)
AMOVA	analysis of molecular variance
ATP	Adenine tri-phosphate
bp	base pairs
CI	confidence interval
Cl	Chloride
CMZ	Cape mountain zebra
DAPSA	program for DNA and protein Sequence analysis (Harley 2001)
D_C	the arc distance of Cavalli-Sforza and Edwards (1967)
D_S	Nei's standard genetic distance (1972)
D_{xy}	average nucleotide divergence
dH_2O	distilled water
DNA	deoxyribonucleic acid
DNTPs	deoxyribonucleoside triphosphates
DRC	Democratic Republic of Congo
EDTA	ethylenediaminetetra-acetic acid
ELB	Erythrocyte lysis buffer
et al.	and company
Exp	expected
Fig.	Figure
Figs	Figures
F_{ST}	fixation index of Wright (1969): the variance of allele frequencies of heterozygotes at a locus
GDNR	Gariep Dam Nature Reserve
H	Hydrogen
H_E	Unbiased expected heterozygosity Nei (1978)
HKY	substitution model (Hasegawa *et al.* 1985)
HMZ	Hartmann's mountain zebra
HWE	Hardy-Weinberg equilibrium
H_O	Observed heterozygosity
I	inertia, principal component goodness-of-fit statistic
IAM	infinite alleles model
IBD	isolation by distance
IUCN	International Union for the Conservation of Nature
K	Potassium
kb	kilobases
km, km^2	kilometre, square kilometre
KNP	Karoo National Park
KNR	Karoo Nature Reserve
M	the ratio of the number of alleles present to the spread of alleles
MET	Namibian Ministry of the Environment and Tourism
mg	milligram
Mg	Magnesium
MHC	major histocompatibility complex
min	minute
mL	millilitre
ML	maximum likelihood
mm	millimetre
mM	millimolar

MP	maximum parsimony
mtDNA	mitochondrial DNA
Mya	million years ago
Myr	million years
MZ	mountain zebra
n	number in sample
Na	Sodium
N_A	the number of alleles
N_e	effective population size
N_em	the effective number of migrants per generation between two populations
N_em_{PA}	N_em estimated by the private alleles method (Slatkin, 1985)
N_em_{RST}	N_em estimated from R_{ST} according to Ciofi and Bruford (1999)
N_F	founder number
ng	nanogram
$(NH_4)_2SO_4$	ammonium sulphate
NJ	neighbour-joining (Saitou and Nei, 1987)
nm	nanometre
NPWS	National Parks and Wildlife Service of Botswana
N_S	the number of animals sampled in each population
NS	not significant
N_P	the number of polymorphic loci in each population
No.	Number
Obs	observed
p, P, p-value	Probability value of rejecting a null hypothesis
p(1)	average private allele frequency
pers. comm.	personal communication
PCA	principal components analysis
PCR	polymerase chain reaction
pH	per Hydrogen
PNK	Polynucleotide kinase
PZ	plains zebra
R^2	coefficient of regression
RAPD	randomly amplified polymorphic DNA
Ref.	Reference
Rho	Goodman's (1997) estimator of R_{ST}
RNA	ribonucleic acid
rpm	revolutions per minute
rRNA	ribosomal RNA
RSA	Republic of South Africa
R_{ST}	fixation index of Slatkin (1995): the proportion of the total variance in allele size that is present between populations
SDS	sodium dodecyl sulphate
S.E.	standard error
sec.	second/seconds
SMM	stepwise mutation model
STE	sodium chloride-Tris-EDTA
T(A)	primer annealing temperature
T	divergence time (in generations), $T = (\delta\mu)^2/2v$
TBE	Tris-boric acid-EDTA
TE	Tris-EDTA
U	Units
v	mutation rate of microsatellite loci
Va	variation
W	Watts

Symbols

α	alpha parameter of gamma distributed substitution rate heterogeneity
$\alpha^{32}P$	phosphorus isotope 32, emitting alpha radiation
°C	degrees Celsius
$(\delta\mu)^2$	distance statistic of Goldstein *et al.* (1995)
μ	mutation rate of mitochondrial DNA
μCi	microCuries
μg	microgram
μL	microlitre
μM	micromolar
%	percent
Φ_{ST}	fixation index analogous to F_{ST} used in AMOVA
π	mean nucleotide diversity
π^1	mean nucleotide diversity calculate by the Kimura 2-parameter method
π^2	mean nucleotide diversity calculated by maximum likelihood
θ	an unbiased estimator of F_{ST}
λ	wavelength
χ^2	Chi-squared

Species and subspecies

Latin name	Common name
Alcelaphus buselaphus	red hartebeest
Crocuta crocuta	spotted hyaena
Connochaetes taurinus	blue wildebeest
Cymbopogon plurinoides	turpentine grass
Damaliscus pygargus	bontebok
Diceros bicornis	black rhinoceros
Enneapogon scoparius	no common name
E. caballus, *Equus caballus*	horse
E. greyvi, *Equus greyvi*	Grevy's zebra
E. quagga, *Equus quagga*	plains or common zebra
E. q. antiquorum, *Equus quagga antiquorum*	Damara zebra
E. q. burchelli, *Equus quagga burchelli*	Burchell's zebra (extinct)
E. q. chapmani, *Equus quagga chapmani*	Chapman's zebra
E. q. selousi, *Equus quagga selousi*	Selous' (Chapman's) zebra
E. q. crawshayi, *Equus quagga crawshayi*	Crawshay's zebra
E. q. granti, *Equus quagga granti*	Grant's zebra
E. q. boehmi, *Equus quagga boehmi*	Boehm's zebra
E. q. quagga, *Equus quagga quagga*	Cape quagga (extinct)
E. q. wahlbergi, *Equus quagga wahlbergi*	Wahlberg's zebra
E. zebra, *Equus zebra*	mountain zebra
E. z. hartmannae, *Equus zebra hartmannae*	Hartmann's mountain zebra
E. z. zebra, *Equus zebra zebra*	Cape mountain zebra
Heteropogon contortus	assegaai grass
Panthera leo	lion
Panthera pardus	leopard
Setaria neglecta	bristle grass
Struthio camelus	ostrich
Themeda triandra	rooigras (red grass)

Table of Contents

Chapter 1: Introduction

Zebras, the most diverse group of wild equids, are distinctively coloured stenoids with black stripes upon white or buff ground colour. This unmistakable striping pattern has become one of the most recognised symbols of the entire African continent, even though zebras are exclusively sub-Saharan in distribution.

Three distinct species are recognised, the mountain zebra, the plains zebra and Grevy's zebra. The mountain zebra (*Equus zebra* Linnaeus, 1758) is the smallest in size and most specialised of the zebras and is restricted to the arid western escarpment and fold mountain systems of southern Africa. Two subspecies are recognised, one of which, the Cape mountain zebra (*E. z. zebra*) of South Africa is threatened, with a total population size of 1600 and fewer than 1000 breeding adults (Castley *et al.* 2002). All extant Cape mountain zebra are descendants of no more than 30 individuals from three relict aboriginal populations. The other subspecies, Hartmann's mountain zebra *(E. z. hartmannae*), is commonly found along Namibia's western escarpment.

The second species is the plains zebra (*Equus quagga* Boddaert, 1785*)*, which is most common and widely distributed, exhibiting high pelage diversity across its range from Sudan to South Africa. Phenotypic diversity in *E. quagga* is clinal (Rau, 1978), therefore confounding attempts at subspecific classification. Two subspecies are still extant in the southern African sub-region: Chapman's zebra (*E. q. chapmani*) of Zimbabwe and the Damara zebra (*E. q. antiquorum*) of Angola, Namibia, Botswana, South Africa and Swaziland.

The third species, Grevy's zebra (*Equus grevyi* Oustalet, 1882) is the largest and most narrowly striped. Having a previously wide distribution encompassing much of Africa and the Middle East, this species' range has been reduced to the arid savannas of northern Kenya and southern Ethiopia. Behaviourally, Grevy's zebra differs from the other two species in that adult males are territorial, defending areas which possess good grazing and water (Kingdon, 1997). The species is endangered, with no more

than 5000 animals existing in the wild (Eliot, 1993). The ranges of the two less common species, *E. zebra* and *E. grevyi*, overlap with the wide distribution of the plains zebra. Areas of sympatry include southern Ethiopia and northern Kenya, where *E. quagga* and *E. grevyi* both occur, sometimes forming mixed herds (Kingdon, 1997), and northern Namibia where *E. quagga* and *E. zebra* associate with each other at water holes. Despite the close proximity of zebra species in areas of sympatry, the incidence of hybridisation in the wild is unconfirmed, and few suspected hybrids have been observed.

1.1 Conservation in fragmented habitats

Habitat destruction associated with an increasing human population has resulted in the fragmentation of populations of all extant zebra species and the disruption of ancient migration and dispersal routes, thus obviating evolutionary processes of drift and selection. The worst affected taxa such as the Cape mountain zebra (*E. z. zebra*), the extinct Burchell's zebra (*E. q. burchelli*) and Cape quagga (*E. q. quagga*) are those that have been most closely associated with humans in southern Africa. Ironically, the mountain zebra is the only species that did not show a major population decline in the last 20 years (Eliot, 1993; Kingdon, 1997; Moehlman, 2002). If habitat destruction and population fragmentation in the rest of the continent continue unabated, some subspecific forms of the now abundant *E. quagga* may become endangered (Kingdon, 1997). The metapopulation management strategy of Lande (1993) and Ballou *et al.* (1995) was designed specifically to aid the conservation of populations in disturbed or fragmented systems. A metapopulation structure consists of independently evolving subpopulations across the entire specific or subspecific range. Implementation of this kind of management strategy entails the reintroduction of populations to areas from which they were previously exterminated, thereby increasing the number of geographically separated demes. This will have the advantage of attenuating the spread of infectious diseases. However, the existence of (usually small) populations in isolation will invariably lead to a decrease in genetic diversity as a result of genetic drift. The rectification of this problem entails sustaining and monitoring large populations and also the intermittent movement of animals between demes in order to maintain genetic diversity across the metapopulation (Lande, 1995; Bowland *et al.* 2001).

Genetics is able to play a large role in conservation as molecular data are objective and are able to determine the structure of populations with high statistical precision.

This information is essential in identifying high risk taxa. Effective conservation of Africa's declining zebra populations will only occur once unequivocal baseline genetic data are obtained and incorporated into management programmes. The initial effect of this will be the establishment of contiguous or semi-contiguous conservation areas that will not only enable gene flow via dispersal to take place, thus maintaining genetic diversity, but will also have the more important effect of protecting whole ecological systems. This will benefit all taxa in the ecosystem, for most of which genetic information is either unknown or difficult to obtain.

1.2 Microsatellites

Microsatellites have become the molecular markers of choice in conservation genetic studies for a wide variety of taxa. They consist of short tandemly repeated (STR) elements, 1-6 base pairs in length, distributed across the eukaryote genome (Tautz and Renz, 1984; Edwards *et al.* 1991). Microsatellites range in size, but generally contain less than 30 repeats and dinucleotide motifs, especially "CA", are the most common in animals (Jarne and Lagoda, 1996). An organism's genotype at a particular microsatellite locus is determined by how many repeats each of its alleles contain. Since each of these alleles is inherited in Mendelian fashion, codominant microsatellite data are robust and amenable to analysis by a wider variety of population genetic parameters than haplotypic and minisatellite data. The repeat nature of microsatellites renders them highly mutable (Henderson and Petes, 1992). During DNA replication, DNA polymerase may lose track of its place, and either leaves out repeat units or adds too many repeat units. The replicated strand therefore has a different number of repeats as the parent strand. The probability of slippage is thus much higher for this type of DNA arrangement (Schloetterer and Tautz, 1992; Jarne and Lagoda, 1996). Most loci are therefore characterised by high levels of polymorphism and this provides one of the most useful attributes of microsatellite data with respect to rare species conservation (Bruford and Wayne, 1993).

However, mutation rates as high as 10^{-3} mutations per generation (Jeffreys *et al.* 1988) to 10^{-4} mutations per generation (Levinson and Gutman, 1987) can also be a drawback, as constraints in allele size and the ensuing homoplasies render microsatellites less useful at higher taxonomic levels (Bowcock *et al.* 1994; Feldman *et al.* 1997; Paetkau *et al.* 1997). Another drawback arises if there is no genetic

information about the organism being studied, or no information even from a closely related species. In such cases, a considerable amount of time must be expended in isolating and cloning microsatellite loci by screening a size-selected genomic library of the organism of interest.

1.2.1 Genetic drift, inbreeding and outbreeding

Genetic drift is a process where independently evolving populations which are exposed to the same selection pressures may diverge owing to a random sample of alleles being passed on to each subsequent generation (Buri, 1956). Even adjacent populations in relative isolation will differentiate independently as a natural consequence of this random component of gene-frequency change (Suzuki et al. 1989). In widely distributed species where large-scale migration is relatively restricted, local genetic drift will result in a general structure of isolation by distance (IBD) of populations across the species distribution (Slatkin, 1993). When a species or group under study has passed through one or more population bottleneck(s), the observed genetic differentiation or loss of diversity owing to genetic drift may increase very quickly (Hedrick, 1999). The duration of the genetic bottleneck will also affect the genetic structure of an affected population. If the duration of the bottleneck is short and the population multiplies rapidly, the degree of genetic loss will be minimised. However, if the genetic bottleneck is severe and lasts for a number of generations, random genetic drift will rapidly drive alleles either to extinction or fixation. In small post-bottlenecked populations, individuals have no choice but to inbreed with close relatives. The chances of lethal or deleterious genes being represented in the homozygous state increases as close siblings are more likely to possess copies of the same lethal or deleterious alleles (Lynch and Gabriel, 1990). Inbreeding depression is therefore any condition which compromises the fitness of an individual as a result of the genetic consequences of breeding with close relatives (Charlesworth and Charlesworth, 1987).

Genetic drift is not restricted to neutral microsatellite loci and it can be assumed that the degree of homozygosity in an inbred population is a reflection of that of the entire genome, except where there is balancing selection. Many recessive alleles in a population are harmful to a greater or lesser degree (Crow, 1986). These alleles whilst rare are effectively hidden to selection and persist in an outbreeding population because they exist only in the heterozygous state, and their effects are masked by the dominant alleles. In an inbred population, where the proportion of homozygous

individuals increases, the deleterious effect of homozygous recessive genotypes becomes realised (Charlesworth and Charlesworth, 1987). A further consideration is when heterozygotes have a superior fitness. In this case (heterosis), inbreeding is associated with decreased vigour, poor survival and reduced fertility as heterozygosity is lost (Crow, 1986). Such lower long-term fitness due to reduced genetic variation has been suggested by O'Brien *et al.* (1985), Roelke *et al.* (1993) and Keller *et al.* (1994), primarily as a consequence of selection against inbred individuals (homozygotes).

With a multiallelic, highly variable marker such as a microsatellite locus, different inbred populations may become randomly fixed for different alleles. Due to a higher initial number of alleles, changes in allele frequencies will develop more frequently in an inbreeding population. The high number of random changes in allele frequencies of two independently inbreeding populations, derived from the same outbred metapopulation, is more likely to result in two very differentiated populations than in two very similar ones. Therefore, if genetic differentiation is to be found in rare or endangered species, it is most likely to be found in microsatellite loci.

In some cases outbreeding depression may occur if two populations, which have diverged and independently accrued evolutionary significant changes over time, are bred together (Waser, 1993). From a conservation perspective, an understanding of outbreeding depression in important, since translocation among natural populations is becoming increasingly common (Edmands, 1999). The reproductive isolatory mechanisms apparent within the Equidae, increase the threat of outbreeding depression in zebras. In cases where translocations across large distances may be necessary, the genetic divergence, even between subspecies, must first be quantified before management decisions can be made.

1.3 Mitochondrial DNA

Mitochondrial lineages offer conservation and population geneticists a wealth of unique molecular information. Mitochondrial DNA (mtDNA) is non-recombining and maternally inherited, thereby allowing for the determination of common ancestry through shared derived characters. It is also present in high copy number in each cell and may therefore persevere for longer in ancient, weathered and faecal substrates. Owing to

more relaxed constraints on DNA repair mechanisms (Cann *et al.* 1984), the rate of mutation (and hence divergence between lineages) of the mtDNA molecule is 5-10 times higher than nuclear DNA in most animals (Ferris *et al.* 1983; DeSalle *et al.* 1986). The relatively rapid accumulation of selectively neutral mutations (primarily base substitutions) makes it possible to uncover phylogenetic relationships even in the lower taxonomic ranks (e.g. between species and subspecies). However, the rate of mutation is not constant throughout the mitochondrial genome (Ballard and Kreitman, 1995), nor is it constant for the same genes between species (Cann *et al.* 1984). This therefore could cause deviation from a strict application of Kimura's (1983) neutral theory of molecular evolution. However, differing rates of evolution for various mitochondrial regions conveniently enables evolutionary resolution to be effected at different historical time scales, provided that the rate of base substitution is comparable across all studied taxa. Substitution rate heterogeneity must therefore be considered and corrected for, especially when a molecular clock is used to date the coalescence of mitochondrial lineages.

The control region is the origin of replication of the mitochondrial DNA molecule. As it is not coding, the control region exhibits the highest sequence variability in the mitochondrial genome. The control region is thus particularly useful in intraspecific phylogeography (Eizirik *et al.* 1998; Matthee and Robinson, 1999; Eizirik *et al.* 2001; Troy *et al.* 2001; Vila *et al.* 2001). Furthermore, haplotype structuring at the level of the control region provides a complement to microsatellite analyses as the overlap in the resolving powers of both markers represents the interface of population and phylogenetics.

The phylogenetic relationships of the zebroid equids from control region data have already been determined (Oakenfull *et al.* 2000) but due to a dearth in sample availability, some southern African subspecies of both *E. zebra* and *E. quagga* have not been well represented and the data were analysed only in a phylogenetic context. The intraspecific structuring of mountain and plains zebra haplotypes within the southern African sub-region is also of conservation interest as genetic support for subspecific partitioning has not been thoroughly examined, especially in the case of subspecificity in mountain zebras.

1.4 General aims

The general objectives of this study are:

- First, to obtain a definitive understanding of the genetic status and structure of the key populations of the two species of zebra (*Equus zebra* and *Equus quagga*) occurring in the southern African sub-region.
- Second, to attempt to explain the evolutionary or other mechanisms likely to have led to the observed genetic scenarios.
- Finally, to use the results from the first two aims in a practical way to assist in the conservation management of southern African zebras.

Chapter 2: Methodology

2.1 Study area

Samples were collected from a wide range of locations throughout the sub-region. Mountain zebras were sampled from 15 locations in Namibia and South Africa. Samples from nine populations of plains zebra, representing the major demes in Namibia, Botswana, Zimbabwe and South Africa were obtained. The geographic situation of each sample locality is represented in Figure 2.1 and described in Table 2.1. The vague nature of geographic data supplied with each sample limited the

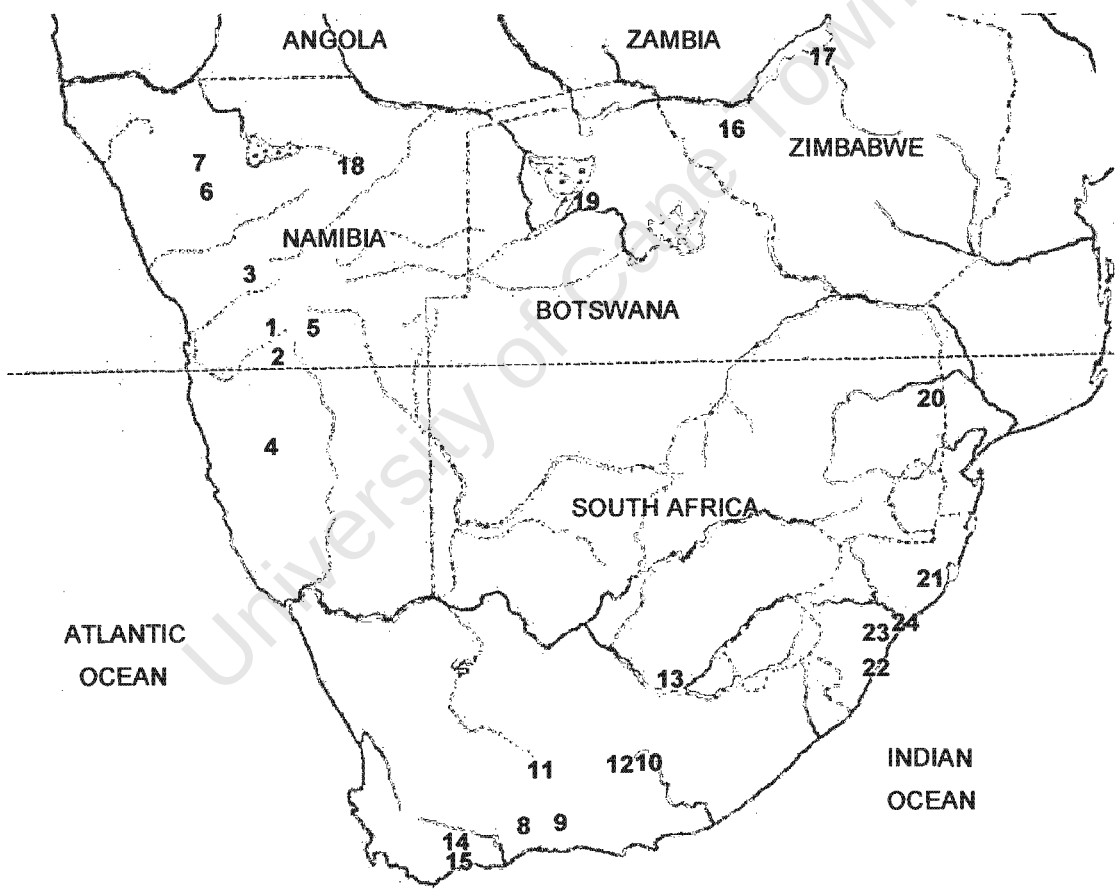

Figure 2.1. Map of the southern African sub-region showing the origin of mountain zebra (*E. zebra*, locations 1-15) and plains zebra (*E. quagga*, locations 16-24) samples. A brief description of each sampling locality is given in Table 2.1.

precision to which population locality could be determined. The various points plotted on Fig 2.1 represent the estimated midpoint of the area from which each sample was actually collected.

Table 2.1. Brief description of 15 mountain zebra and nine plains zebra sample localities. Reference numbers (Ref. no.) correspond to those mapped on Fig. 2.1. Words in bold are the names by which samples will hitherto be referred. KH, Khomas Hochland; RSA, Republic of South Africa

Mountain zebra		Plains zebra	
Ref. no. (Fig 2.1)	Sample locality	Ref. No. (Fig 2.1)	Sample locality
1	**Khomas Hochland**, mid-Central region, Namibia	16	**Western Zimbabwe**, Matetsi, Deka and surrounding Safari Areas
2	**Gamsberg**, adjoining KH, mid-Central region, Namibia	17	**Zambezi Valley**, Charara, Chete and surrounding Safari Areas
3	**Erongo** Mountains, north-Central region, Namibia	18	**Namibia**, private game farms, east of Etosha National Park
4	**Naukluft** Mountains, south-Central region, Namibia	19	**Botswana**, Ngamiland and Chobe Provinces, northern Botswana.
5	**Auasberg**, extralimital, introduced population	20	**Lowveld**, private game farms, erstwhile Transvaal, RSA
6	**Kamanjab District**, Damaraland, Northern region, Namibia	21	**Hluhluwe-Umfolozi** Park, Mfolozi Valley, KwaZulu, RSA
7	**Otjivasandu**, Etosha National Park, Northern region, Namibia	22	**Vernon Crookes** Nature Reserve, Umzinto, KwaZulu, RSA
8	**Gamka** Mountain Nature Reserve, Klein Karoo, RSA	23	**Albert Falls** Nature Reserve, Midlands, KwaZulu, RSA
9	**Kammanassie** Nature Reserve, Klein Karoo, Uniondale, RSA	24	**Harold Johnson** Nature Reserve, Tugela Mouth, KwaZulu, RSA
10	**Mountain Zebra National Park**, Bankberg, Cradock District, RSA		
11	**Karoo National Park**, Nuweveldberg, RSA		
12	**Karoo Nature Reserve**, Sneeuberg, Graaff-Rienet, RSA		
13	**Gariep Dam Nature Reserve**, Gariep River, Bethulie, RSA		
14	**Bontebok National Park**, Overberg, RSA		
15	**DeHoop Nature Reserve**, Overberg, RSA		

2.2 Collection of specimens

As the expense of sampling large mammals may sometimes be prohibitive, a wide variety of collection techniques was employed to obtain the 479 zebra samples used in this study. A number of different sources were used: fresh blood, fresh tissue, dried skin, dried tissue and faeces. Details of these methods as well as the various sampling locations for each of the two zebra species studied are described below. And unless otherwise stated, all game capture was conducted with the aid of a helicopter, provided by various agencies.

2.2.1 Mountain zebra (*Equus zebra*)

2.2.1.1 Cape mountain zebra (*E. z. zebra*)

Gamka

Due to the vulnerability of the small aboriginal population at Gamka Mountain Nature Reserve, invasive sampling was not permitted. Therefore samples in the form of pieces of tissue, tendon and skin, weathered for up to 10 years (n=10) and faeces (n=8) were collected from the field. Samples from two tanned skins, one from Mr T. Barry of Gamka Mountain Nature Reserve and the other from the South African Museum, were also taken.

Kammanassie

A single weathered sample was collected from Kammanassie Nature Reserve along with three faecal samples. Dr P. Morkel, the veterinarian heading the South African National Parks Game Capture Unit, collected four EDTA-treated whole blood samples during a game capture operation and Ms G. Cleaver of Kammanassie Nature Reserve collected skin from the remains of four animals that had died in a variety of mishaps while moving between watering points and feeding grounds.

DeHoop

Four skin samples were obtained using the biopsy darting technique (Karesh *et al.* 1987). A further 13 samples were taken from dry salted skins that were stored at DeHoop Nature Reserve. These animals had died at various times during the last 10 years.

Cradock-derived populations

Samples from the Mountain Zebra National Park were obtained by biopsy darting (n=8), from tanned skins held by the South African Museum (n=5) and from a single weathered skin sample of a dead animal collected by Mr J. de Klerk of South African National Parks.

Dr P. Meyer, also of the South African National Parks Game Capture Unit, collected EDTA-treated whole blood samples from 10 Cape mountain zebra at the Karoo National Park (KNP) as well as ear biopsies from a further three transportation fatalities from the same Park.

Mr H. Barnard immobilised 12 Cape mountain zebra at the Karoo Nature Reserve (KNR) near Graaff-Reinet in the Eastern Cape Province of South Africa. Blood samples were taken from all animals by venipuncture using EDTA-treated tubes, and a dried skin sample was collected from a road kill with the help of Mr P. Burdett.

Fourteen EDTA-treated whole blood samples were collected by Dr P. Morkel from the entire Cape mountain zebra population at Bontebok National Park (BNP), near Swellendam.

Dr P. Nel of Free State Nature Conservation provided 10 EDTA-treated whole blood samples of Cape mountain zebra from the Gariep Dam Nature Reserve (GDNR).

2.2.1.2 Hartmann's mountain zebra (*E. z. hartmannae*)

All samples taken were in the form of dry-salted skin pieces, obtained from a variety of Namibian tanneries, taxidermists and skin dealers. Whole skins were identified by coded tags and from these the exact farm from which each skin originated could be traced.

Central region

Skin samples from populations in Central Namibia were collected at the Western Trading Corporation, Windhoek (Khomas Hochland, n=23; Gamsberg, n=11; Naukluft, n=2), the Nakara Tannery, Windhoek (Khomas Hochland, n=24; Erongo, n=2; Auasberg, n=5; Naukluft, n=1) and Trophaendienste Taxidermy, Windhoek (Khomas Hochland, n=8; Erongo, n=3; Naukluft, n=6).

Northern region

The Kunene Province of Namibia, here referred to as the Northern region to contrast it with the Central region which spans four provinces, is inhabited by a large number (n=6730) of free ranging mountain zebras (Hack *et al.* 2002). The bulk of all skin samples from the Northern region originated from the Otjivasandu district (n=101) on the western border of Etosha National Park. These skins were available as the result of a recent cull by the Namibian Ministry of the Environment and Tourism (MET). Other sources for this region include the Nakara Tannery, Windhoek (Kamanjab, n=7) and Otjiwarongo Taxidermy, Otjiwarongo (Kamanjab, n=8),

2.2.2 Plains zebra (*Equus quagga*)

Samples in the form of either dry-salted pieces of skin or EDTA-treated whole blood were obtained from a variety of sources across the region south of the Zambezi River.

2.2.2.1 Namibia

Namibian EDTA-treated whole blood samples of Damara zebra (*E. q. antiquorum*) from Etosha National Park (n=3) were collected by Mr R. Rau of the South African Museum. Dry-salted skin samples from private game reserves to the east of Etosha National Park were obtained from the Nakara Tannery, Windhoek (n=16) and from Trophaendienste Taxidermy, Windhoek (n=4).

2.2.2.2 Botswana

Ms D. Peake of Mochaba Safari Services, Maun, provided 15 dry-salted skin pieces from plains zebra of the Ngamiland and Chobe Provinces of northern Botswana.

2.2.2.3 Zimbabwe

Chapman's zebra (*E. q. chapmanni*) skin samples (Western Zimbabwe, n=39; Zambezi Valley, n=8) were obtained from Ms T. Swann of Bromley Game Skin Tannery, Harare.

2.2.2.4 Lowveld

Ms W. da Costa of Trans African Taxidermists, Randburg, South Africa, provided 27 dry-salted skin samples from eight different private game farms from the erstwhile

Transvaal Province of South Africa as well as a single sample from an unknown location in Tanzania. A further seven skin samples, representing two separate game farms were obtained from Oryx Taxidermy, Krugersdorp, South Africa. As almost all privately owned reserves in this area were originally seeded with Kruger National Park stock, the specimens obtained from these reserves may be regarded as a representative sample of the large population presently inhabiting the South African Lowveld. As it is impossible to determine the exact area(s) in Lowveld from which the animals originated, the north-south midpoint of the Kruger National Park (Location 20, Fig 2.1, page 7) was estimated as the geographic origin of Lowveld samples.

2.2.2.5 KwaZulu

All Damara zebra samples obtained from KwaZulu were in the form of EDTA-treated whole blood and were taken as part of another study by Bowland *et al.* (2001) that investigated allozyme and RAPD variation in seeded and natural populations of plains zebra. The 66 samples were provided by Dr P. Taylor of Durban Natural Science Museum (Hluhluwe-Umfolozi Park, n=23; the Vernon Crookes Nature Reserve, n=19; the Albert Falls Nature Reserve, n=17 and the Harold Johnson Nature Reserve, n=7).

2.3 DNA extraction

2.3.1 Standard procedure

The SDS-Proteinase K/phenol-chloroform protocol, one of the most widely used methods of DNA extraction (Sambrook *et al.* 1989), was used to isolate DNA from the bulk of all samples obtained. First, SDS is used to disrupt the phospholipid membranes of cells while Proteinase K indiscriminately digests proteins, releasing DNA from the nucleus. Subsequent phenol extractions separate the DNA-containing lysate from proteinacious material and a chloroform/isoamyl alcohol (24:1) extraction continues this process whilst also removing phenol residues.

Four volumes of erythrocyte lysis buffer (ELB: 0.32 M sucrose, 10 mM Tris (pH 7.6), 5 mM $MgCl_2$, 1% [v/v] Triton X 100) were added to 4 - 10 mL of whole blood, collected in EDTA tubes. The mixture was centrifuged in a 50 mL polypropylene tube for 5 min at 3000 rpm in a Sigma 302K centrifuge. The supernatant was carefully discarded and the lymphocyte pellet re-suspended in two volumes of ELB, at which time the centrifugation step was repeated. The supernatant was again discarded and the pellet re-suspended

in 700 µL sodium chloride-Tris-EDTA (STE) isotonic lysis buffer (0.15 M NaCl, 1 mM EDTA, 50 mM Tris (pH 8.0)) and transferred to a 1.5 mL microcentrifuge tube. In the case of fresh (muscle) tissue, no more than 50 mg was added to 700 µL STE lysis buffer. Approximately 100 mg of salted or tanned skin or weathered tissue or skin was washed three times in double-distilled water (dH_2O) for five hours each, prior to transfer to 700 µL STE buffer. For faecal samples, the surface layer believed to contain sloughed cells from the rectal mucosa, was scraped off with a sterile scalpel blade and added directly to 700 µL STE buffer.

Sodium dodecyl sulphate (SDS) was added to a final concentration of 1% [w/v] and Proteinase K (Boehringer Mannheim) to 0.2 mg/mL. The reaction was incubated overnight at 55 °C and in the case of weathered samples, tanned skins and faeces, the lysis procedure was prolonged. The Proteinase K concentration was increased to 0.5 mg/mL for tanned museum skins, faeces and weathered material.

DNA was extracted from the lysate twice by phenol/chloroform/isoamyl alcohol in the proportion of 25:24:1 and then once by chloroform/isoamyl alcohol (24:1). Both phenol and chloroform are effective protein denaturants with differing properties (Ausubel *et al.* 1987). Isoamyl alcohol was used to aid the separation of organic and aqueous phases whilst the denatured proteins accumulated in the interface between the phases during centrifugation at 14 000 rpm in a Sigma 2MK bench top microcentrifuge. DNA was isolated in the aqueous phase and this was carefully removed after each step. The extracted DNA was precipitated with the addition of 0.1 volumes 3 M sodium acetate and 1 volume of isopropanol. The precipitated DNA was centrifuged at 14 000 rpm in a refrigerated Sigma 2MK bench top microcentrifuge set to 0 °C. The pellet was washed in 1 mL ice-cold (-20 °C) 70% ethanol. Washing the DNA with salt soluble 70% ethanol effectively desalts the DNA while also removing traces of isopropanol. The DNA was air dried for 30 min before being transferred to between 50 and 500 µL Tris-EDTA (TE) buffer (10 mM Tris (pH 8), 1 mM EDTA (pH 8)), in which it was left to dissolve overnight at 55 °C.

2.3.2 Faecal samples

Since blood and fresh tissue samples from the most vulnerable mountain zebra populations were very difficult to obtain, a concentrated effort was made to obtain amplifiable DNA from faeces. Apart from the extraction procedures above, various

methods of faecal preservation were tested: fresh faeces were collected from an accessible plains zebra population and dried at 70 °C, at room temperature, at 4 °C and in full sunlight. All these, as well as a fresh faecal sample, were subjected to all the above described methods of DNA extraction. DNA from faeces was also isolated using potato flour adsorption (Deuter *et al.* 1995), silica adsorption (Taberlet *et al.* 1997) and by use of the Qaiamp DNA extraction kit (Qaigen).

2.3.3 DNA quantification

The isolate quality and yield were assayed by ultra-violet spectrophotometry. Optical densities at 260 and 280 nm allow the quantification of DNA using the equation: Quantity (μg/μL) = OD(260) x 50 x D, where OD(260) is the optical density or absorbance of electromagnetic radiation at λ = 260 nm and D the factor by which the tested sample is diluted. Quality of isolated DNA was determined by calculating the ratio between optical densities at 260 nm and 280 nm. For pure DNA isolates, the calculated ratio should ideally fall between 1.8 and 2 (Sambrook *et al.* 1989).

2.4 Microsatellites

2.4.1 Optimisation

Primers that amplify microsatellite loci in horses are widely available because of the high commercial value of unequivocal paternity determination. However, horse primer pairs may not always amplify microsatellites in other equids. Primer pairs that amplify uninterrupted dinucleotide repeated elements of 20 microsatellite loci were tested on plains and mountain zebras (see Table 2.2). As interrupted repeats tend to be less mutable (Jarne and Lagoda, 1996), only uninterrupted horse repeat motifs were chosen to allow comparison of all the resulting data without the statistical complications of differences in mutation rate between loci. In most cases, these loci contained the highest amount of variation in horses.

2.4.2 Primer end-labelling

Microsatellite loci were visualised by autoradiography. Primers were obtained from the University of Cape Town Oligonucleotide Synthesising Unit and diluted to 20 μM in dH_2O. Forward primer end-labelling was performed in 1 x Polynucleotide kinase (PNK)

reaction buffer (70 mM Tris-HCl,10 mM $MgCl_2$, 5 mM dithiothreitol pH 7.6; Bioline) with 0.08 μM forward primer, 2 μCi/μL α^{32}P ATP (AEC Amersham) and 0.5 U/μL T4 PNK (Bioline). PNK used the energy of the ATP molecule to attach the ^{32}P isotope to the 5' end of the forward primer. The reaction was allowed to incubate for 90 min and then stopped by denaturation of PNK at 80 °C for 2 min.

2.4.3 Amplification

PCR was carried out in 1 x Polymerase reaction buffer (16 mM $(NH_4)_2SO_4$, 67 mM Tris-HCl (pH 8.8), 0.01% Tween-20; Bioline), 3 mM $MgCl_2$, 0.25 mM DNTPs (AEC Amersham), 0.08 μM end-labelled forward primer, 0.08 μM reverse primer, 0.05 U/μL Taq Polymerase (Bioline) and 20 ng template genomic DNA in a total reaction volume of 10 μL.

2.4.4 Thermocycling

Reactions were cycled on a Stratagene Robocycler according to the following protocol: initial denaturation for 2 min at 94 °C, followed by 35 cycles of denaturation at 94 °C for 45 sec, annealing at the prescribed annealing temperature (Table 2.2) for 45 sec and extension at 72 °C for 45 sec. A final extension at 72 °C for 10 min terminated thermocycling.

2.4.5 Weathered and faecal DNA samples

The same general protocol was followed but $MgCl_2$ concentration was increased to up to 4 mM, DNTPs up to 0.5 mM and the total reaction volume to 20 μL. Positive and negative controls were used with all sets of reactions performed. Denaturation during thermocycling was carried out at 96 °C. A range of annealing temperatures was used to optimise PCR products for each primer pair and the final extension step increased to 30 min.

2.4.6 Electrophoresis

Each 10 μL (or 20 μL) microsatellite PCR reaction was stopped by the addition of 8 μL (or 16 μL) PCR stop solution (98% deionized formamide, 10 mM EDTA (pH 8.0), 0.025% xylene cyanol FF, 0.025% bromophenol blue). Prior to electrophoresis, the

samples were denatured at 96 °C and cooled on ice. Electrophoresis was carried out in 1 x TBE electrophoresis buffer (89.15 mM Tris, 88.95 mM boric acid, 2 mM EDTA (pH8.0)) using a Hoefer Poker Face II SE 1600 vertical electrophoresis apparatus. Up to 60 reactions (lanes 3.5 mm wide) were loaded onto a 0.4 mm thick, 432 mm long 6% polyacrylamide gel (6% 20:1 acrylamide:bis-acrylamide solution, 8M urea, 1 x TBE electrophoresis buffer) and electrophoresed at a constant power of 65 W for 2-3 hours. The A-T sequence of the M13 genome downstream of the -40 primer annealing site was loaded in two lanes among the samples for subsequent use as a size standard to score alleles. Post-elecrophoresis, the gel apparatus was dismantled and the gel transferred onto Whatman 3MM filter paper, dried at 80 °C for 2 hours, exposed to X-ray film (Agfa or Cronex) in a light-proof cassette for 3 hours to 7 days and then developed.

2.4.7 Optimisation results

Results of the optimisation procedure are applicable to both plains and mountain zebras and will therefore be discussed here. Of the 20 prospective horse microsatellite loci tested, 15 amplified polymorphic fragments in both plains and mountain zebra (Table 2.2). For loci HTG 8 and HTG 10, no amplification product was obtained, regardless of reagent, primer and template titrations. Loci HTG 6 (Fig 2.2) and AHT 16 (Fig 2.3) amplified well but were not polymorphic in either of the two zebra species studied. Locus HMB 5 could be amplified only when DNA was of the highest quality (A260/A280 > 1.9). As many of the DNA samples used in this study were extracted from unfavourable sources such as tanned skins and rotted carcasses, locus HMB 5 was excluded from further use owing to its erratic nature. All other loci amplified well under prescribed or modified conditions.

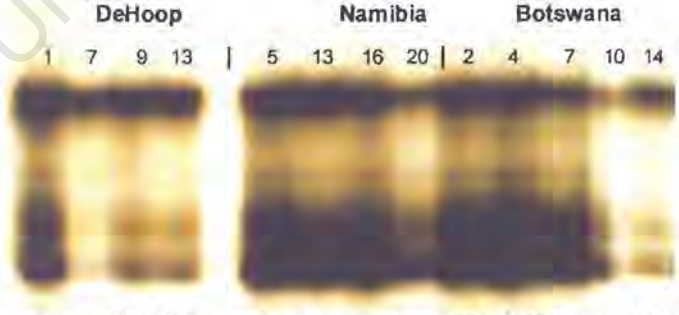

Figure 2.2. Monomorphic PCR products at Locus HTG 6 (Ellegren *et al.* 1992). A selection of individuals of both mountain (DeHoop) and plains zebra (Namibia and Botswana) from three genetically diverse populations was chosen

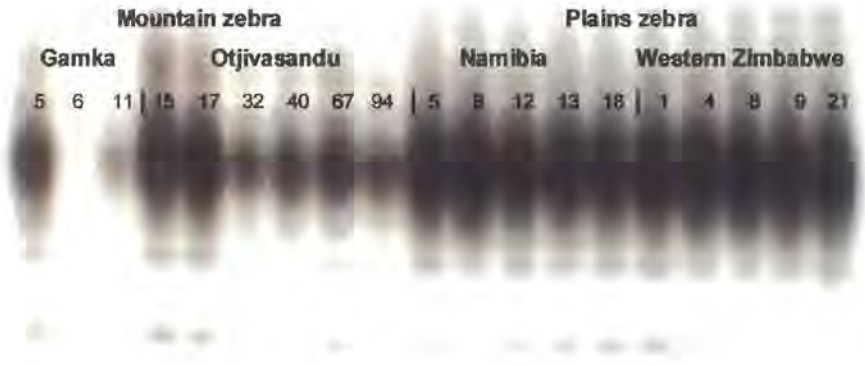

Figure 2.3. Monomorphic PCR products of a selection of two zebra species from four populations at Locus AHT 16 (Swinburne *et al.* 1997). Populations Otjivasandu, Namibia and Western Zimbabwe are free ranging, outbred and usually contain high genetic diversity.

Table 2.2. Summary information for horse (*Equus caballus*) dinucleotide microsatellite loci tested in plains and mountain zebras. T(A) is the prescribed annealing temperature for each locus. The AMP column indicates whether the locus amplified (+) or did not amplify (-) in zebras and loci were denoted as polymorphic (+) or non-polymorphic (-) loci in mountain (MZ) and plains (PZ) zebras.

Locus	Repeat Motif	T(A)	Reference	AMP	Polymorphism	
					MZ	PZ
HMB 1	$(GT)_{18}$	63 °C	Binns *et al.* 1995	+	+	+
HMB 5	$(GT)_{14}$	65 °C	,, ,, ,,	-/+	+	+
HTG 3	$(TG)_{16}$	63 °C	Ellegren *et al.* 1992	+	+	+
HTG 5	$(TG)_{15}$	63 °C	,, ,, ,,	+	+	+
HTG 6	$(TG)_{20}$	55 °C	,, ,, ,,	+	-	-
HTG 7	$(TG)_{19}$	55 °C	Marklund *et al.* 1994	+	+	+
HTG 8	$(TG)_{17}$	55 °C	,, ,, ,,	-	-	-
HTG 9	$(TG)_{17}$	55 °C	,, ,, ,,	+	+	+
HTG 10	$(TG)_{16}$	55 °C	,, ,, ,,	-	-	-
HTG 11	$(GT)_{15}$	55 °C	,, ,, ,,	+	+	+
HTG 14	$(TG)_{14}$	55 °C	,, ,, ,,	+	+	+
HTG 15	$(TG)_{14}$	55 °C	,, ,, ,,	+	+	+
LEX 20	$(TG)_{19}$	55 °C	Coogle *et al.* 1996	+	+	+
LEX 52	$(TG)_{13}$	55 °C	Coogle and Bailey, 1997	+	+	+
VHL 47	$(CA)_{13}$	55 °C	van Haeringen *et al.* 1998	+	+	+
AHT 16	$(TG)_{23}$	60 °C	Swinburne *et al.* 1997	+	-	-
AHT 21	$(GT)_{21}$	60 °C	,, ,, ,,	+	+	+
UCDEQ 505	$(GT)_{17}$	55 °C	Eggleston-Stott *et al.* 1997	+	+	+
TKY 273	$(CA)_{22}$	55 °C	Tozaki *et al.* 2000	+	+	+
COR 014	$(GT)_{21}$	58 °C	Hopman *et al.* 1999	+	+	+

Multiplex PCR, where more than one locus is amplified in the same reaction, was most successful when used to simultaneously amplify two loci (Fig. 2.4). When three loci were multiplexed (Fig 2.5), increased amplification of non-specific by-product and the relative proximity of loci on electrophoresis made scoring of alleles difficult.

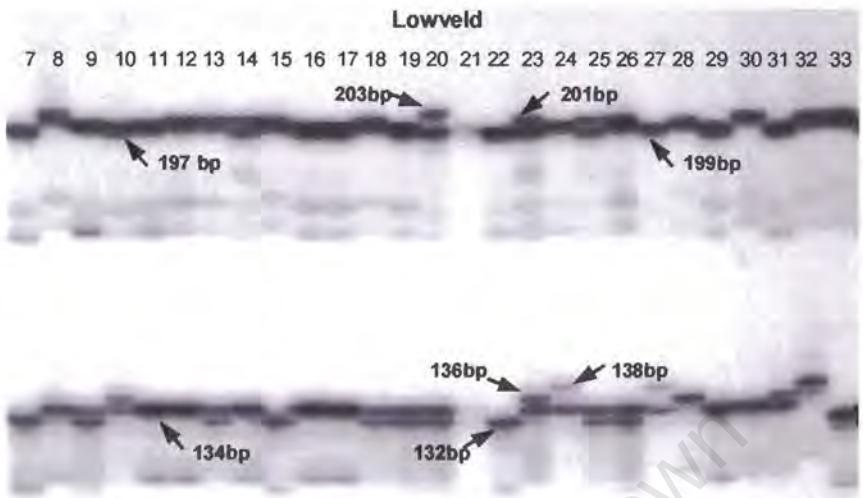

Figure 2.4. PCR amplification products of a two-locus multiplex reaction in a population of plains zebra. Upper locus, LEX 52 (Coogle and Bailey, 1997). Lower locus, VHL 21 (van Haeringen *et al.* 1998).

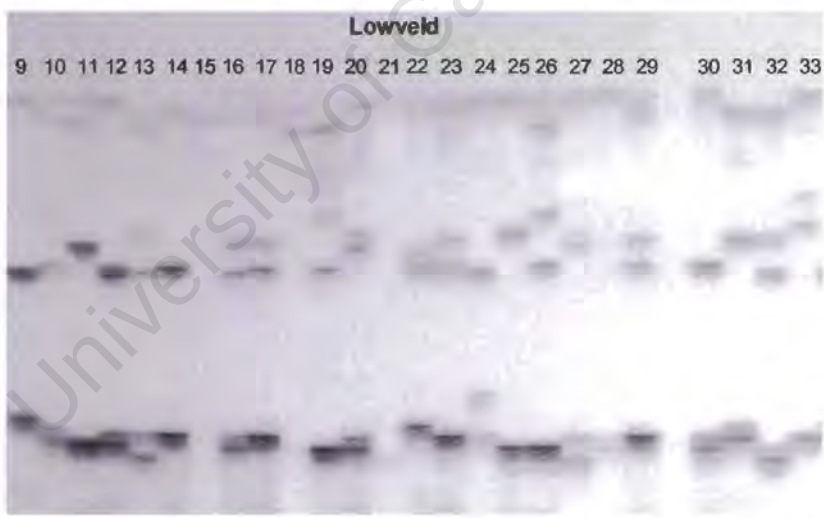

Figure 2.5. PCR amplification products of a three-locus multiplex reaction in a population of plains zebra showing the potential difficulty of scoring alleles when allelic size ranges of loci may overlap. This is exacerbated by individuals that are homozygous for one locus and heterozygous for the other (eg. individuals 19 and 33) and by non-specific amplification products. Upper locus, AHT21 (Swinburne *et al.* 1997). Middle locus, TKY273 (Tozaki *et al.* 2000). Lower locus, HTG15 (Marklund *et al.* 1994).

2.4.8 Microsatellite sequencing

Representative homozygotes of both mountain zebra and plains zebra were chosen for all microsatellite loci used. These loci were then amplified as described in Section 2.4.3, but without end-labelling the forward primer and in a final reaction volume of 50 μL. This large volume was used so that an adequate amount of microsatellite product would be available for sequencing after PCR purification.

The reactions were freed of DNTPs, primers and salts by agarose gel purification. All 50 μL of each reaction were loaded into an 8 mm thick 1% agarose gel, immersed in 1 x TAE electrophoresis buffer (40 mM Tris-acetate, 1 mM EDTA) in a Hoefer HE33 horizontal electrophoresis apparatus. The samples were electrophoresed at a constant potential difference of 3 – 4 volts/cm for 40 – 60 min in the presence of the intercalating dye, ethidium bromide at a concentration of 10 μg/ml. Each PCR product was detected by ultra-violet flourescence on a Spectroline TC-312A ultra-violet transilluminator. A clean scalpel blade was used to excise each microsatellite fragment from the agarose. These were purified using the Qaiex II Gel purification Kit (Qaigen).

Microsatellites were sequenced in the forward and reverse direction using the same primers as were used for the initial amplification. Sequencing was carried out using the Big Dye 2 Sequencing Kit (PE Biosystems) in accordance with the manufacturer's protocol. Sequencing products were cleaned and electrophoresed on either an ABI 377 or ABI Prism automated DNA sequencer at the DNA Sequencing Facility at the University of Stellenbosch.

Non-specific amplification resulted for loci HTG 5, HTG 11 and UCDEQ 505. Exacerbating this was the small size (105 bp) of the largest homozygote at locus HTG 5. While non-specific amplification was usually not problematic when products were radio-labelled and electrophoresed through a high resolution PAGE, the fragments excised from agarose gels contained too many non-specific amplification products, thus nullifying subsequent sequencing reactions.

2.5 Microsatellite analytical techniques and software

Each microsatellite autoradiograph was scored twice independently and the raw data logged on to a Microsoft Excel 97 worksheet. From this, ecumenicism was performed by the program AGaR$_{ST}$ 2.8 (Harley, 2002), enabling direct implementation of the data in a wide range of population genetics software.

2.5.0.1 Linkage disequilibria

Prior to population genetic analysis, the data should be checked for possible linkage disequilibria between loci. If *bona fide* linkage disequilibria exist, the results will be confounded, as alleles of the linked loci will reflect the same patterns of descent. The chance of association between pairs of loci may be tested by means of a Markov chain algorithm (Guo and Thompson, 1992) which tests the null hypothesis that all loci are independent. It should be noted, however, that the statistical association between the allele frequencies of two loci may be due to chance and have nothing to do with linkage (Weir, 1996). The test is an analogue of Fisher's exact test (Fisher, 1935), where the data are reduced to contingency tables of allele frequencies for all pairs of loci in each population. The computation of the actual probabilities of each contingency table is intensive for multiallelic data and is negated by the Markov chain approach. The Markov chain explores the space of all possible contingency tables, finding those tables with the same marginal allele counts as are present in the observed table and then comparing the probability of each of these to that of the observed table. As the Markov chain always starts from the observed sample, a user determined dememorisation phase enables the chain to "forget" its initial state and to make an independent probability comparison. In this way, the p-value is the proportion of tables with a probability less than or equal to the observed table (Schneider *et al.* 2000). The Markov chain is divided into batches to allow for the determination of a standard error. The length of the Markov chain is thus the product of the number of batches and the number of iterations. An overall p-value for each pair of loci across all populations is obtained by the Fisher exact method, using a chi-square test with the number of degrees of freedom equal to two times the number of non-monomorphic populations used in the test. The p-value can be thought of as the probability of being wrong if the null hypothesis is rejected. The benefit of the exact test is that it is non-parametric; it assumes neither random mating nor Hardy-Weinberg equilibrium. Exact tests for linkage disequilibria were implemented by GENEPOP, 3.1c (Raymond and Rousset,

1995a) using 100 batches of 1000 iterations (Markov chain length = 100 000 steps) interspersed with dememorisation phases of 1000 steps.

2.5.1 Intra-population genetic parameters

Allele frequencies were computed by AGaR$_{ST}$ 2.8 (Harley, 2002) and then plotted as allele frequency histograms for each group of mountain zebra populations to represent graphically the spread of alleles at a locus.

2.5.1.1 Genetic variation

Allelic diversity (*A*) and observed heterozygosity (H_O) are two of the most sensitive measures of genetic diversity, especially when populations undergo marked reduction in numbers (Hartl and Pucek, 1994; Spencer *et al.* 2000). Unbiased expected heterozygosity (H_E) was calculated according to Nei (1978) by the program Genetix 4.02 (Belkhir, 2001). Observed heterozygosity (H_O) and allelic diversity (*A*) were computed by the program AGaR$_{ST}$ 2.8 (Harley, 2002). To correct for differing sample sizes, allelic diversity was re-sampled by the jack-knifing technique, using 1000 replicates with the jack-knife sample size equivalent to that of the smallest population. In this instance, jack-knifing is preferable to the bootstrap procedure suggested by Valsecchi *et al.* (1997). If the size of the smallest population is much smaller than the others, each bootstrap replicate sample will be equally small but, with replacement, the chance of missing rare alleles in the sample will increase. Also, as the absolute number of alleles in the smallest population is known, it seems counter-intuitive to use a replacement procedure which will invariably return less than the absolute number of alleles for the smallest population. The means of these parameters for different species and subspecies were compared statistically using a Student's two sample t-test, assuming a normal distribution.

2.5.1.2 Hardy-Weinberg equilibrium

Populations are said to be in Hardy-Weinberg equilibrium (HWE) when genotype frequencies are the product of allele frequencies (Weir, 1996) and both conform to the binomial expansion $p^2 + 2pq + q^2 = 1$, where p and q are allele frequencies in a two-allele system. Natural, free ranging and outbred populations are expected to be in HWE. Deviations from HWE indicate the individuals sampled are not representative of

a randomly mating single population, owing to assortative mating, inbreeding, population structuring (Wahlund effect), selection between the formation of zygotes and the time of sampling, or genotypes having different likelihoods of being included in the sample. Tests for HWE compare the difference in heterozygosity that is observed in a population (H_O) and the heterozygosity that is expected, given the frequency of alleles (H_E). As with the test for linkage disequilibria, a Markov chain simulation was implemented, resulting in a p-value or the probability of rejecting the null hypothesis, that is, that a locus is in HWE. Exact global population and global locus tests for Hardy-Weinberg were performed by GENEPOP 3.1c (Raymond and Rousset, 1995a) using a Markov chain with 100 000 steps and dememorisation phase of 1000 steps. More detailed exact tests for HWE for each population at each locus were computed by Arlequin 2.000 (Schneider *et al.* 2000) using the above set of Markov chain parameters. As each population was subjected to 15 population-locus tests, a sequential Bonferroni test, implemented according to Rice (1989), was used to correct for the chance of falsely rejecting the null hypothesis.

2.5.2 Between-population structuring

2.5.2.1 AMOVA

Population genetic structure can be inferred by means of an analysis of molecular variance (AMOVA). This technique builds on the hierarchical method of analysis of variance by analysing the variance of gene frequencies. The method proposed by Excoffier *et al.* (1992) builds further upon this framework by taking mutation into account. AMOVA tests a genetic hypothesis defined *a priori*, by partitioning the variance component of the data set into hierarchical covariance components, beginning at the intra-individual level. These covariance components are used to generate the fixation index Φ_{ST} for each hierarchical level. For microsatellite data, Φ_{ST} is essentially equivalent to Weir and Cockerham's (1984) θ. Estimation of the genetic structure of a species or subspecies involves the level that describes the variation distributed among defined groups. The AMOVA approach may be applied to both genotypic allelic data and haplotypic DNA sequence data, making it a versatile tool in population genetics. AMOVA was performed by Arlequin 2.000 (Schneider *et al.* 2000). Grouping scenarios for each zebra species are defined in their respective Chapters. The significance of Φ_{ST} was tested using non-parametric permutation tests, which do not assume normality or equality of variance among populations or groups.

2.5.2.2 Exact test of population differentiation

Raymond and Rousset (1995b) proposed that the Markov chain algorithm of Guo and Thompson (1992) be also implemented in an exact test of population differentiation. The GENEPOP 3.1c package (Raymond and Rousset, 1995a) uses the exact test to examine pair-wise population heterogeneity. Contingency tables of columns (populations) and rows (alleles) were tested for heterogeneity, where the independence of rows and columns corresponded to the absence of population differentiation (Raymond and Rousset, 1995b). Probability (p) values indicate the probability of rejecting the null hypothesis of no population differentiation. Exact tests were performed using a Markov chain with 100 000 steps and a dememorisation phase of 1000 steps.

2.5.2.3 Fixation indices

While exact tests of differentiation may determine whether or not significant genetic heterogeneity occurs between pair-wise comparisons, the magnitude of this heterogeneity may be assessed by multilocus fixation indices and genetic distances. Most commonly used is F_{ST} (Wright, 1969), the variance of allele frequencies of heterozygotes at a locus (Avise, 1994). It is computed as $F_{ST} = \sigma_p^2/p(1-p)$, where σ_p^2 is the variance of allele frequencies among populations and p is the observed mean allele frequency. In this way, F_{ST} is standardised relative to the maximum possible allele frequency value. Wright's (1969) original version of F_{ST} is biased by unequal sample sizes. Weir and Cockerham (1984) developed θ, an unbiased estimator of Wright's (1969) F_{ST}, calculated as $\theta = \sigma_a^2/\sigma_a^2 + \sigma_b^2 + \sigma_w^2$, where σ_a^2 is the variance in allele frequency among samples, σ_b^2 is the variance in allele frequency between individuals within a sample and σ_w^2 is the variance in allele frequency within individuals. If samples sizes are equal, $\theta = F_{ST}$. Pair-wise values for Weir and Cockerham's (1984) F_{ST} estimator, θ, were calculated using Genetix 4.02 (Belkhir, 2001) and tested for significance from zero with 10 000 permutations.

F_{ST} and its derivatives are calculated on the assumption that every new mutation within a microsatellite locus results in a new allele, the infinite alleles model (IAM) of Ohta and Kimura (1973). However, it is more plausible that microsatellite loci evolve via the stepwise mutation model (SMM; Kimura and Crow, 1964) where mutations involve the gain or loss of a repeat unit. This model would be characterised by a high proportion of

homoplasic allele classes, that is, alleles of the same size are not necessarily derived from a common ancestor. Slatkin (1995) derived R_{ST} in an attempt to account for these back mutations and to incorporate the SMM of microsatellite evolution. Unlike the analogous F_{ST}, which is derived from allele frequencies, R_{ST} is obtained from the ratio of variances in allele size among populations. More specifically, it is that proportion of the total variance in allele size that is present between populations, calculated as R_{ST} = (SI-SW)/SI, where SI is twice the estimated variance in allele size across populations and SW is twice the estimated variance in allele size within each population (Slatkin, 1995). The drawback of R_{ST} is that, by taking mutation into account, it has a larger variance than F_{ST}.

However, Slatkin's R_{ST} assumes that the variances of all loci are equivalent and that population sample sizes are equal (Goodman, 1997). Calculating R_{ST} over multiple loci involves an initial averaging of variance components over all loci. Consequently, loci with lower variances will contribute less to the final result. To make different loci directly comparable, Goodman (1997) introduced a standardisation protocol to transform the raw data into standard deviations of one from a global mean of zero. Valsecchi *et al.* (1997) showed that Slatkin's R_{ST} was markedly affected by differences in sample size. To correct for this, Goodman (1997) developed *Rho*. This statistic obtains the actual between-population variance component (Sb) by subtracting, from the observed between-population variance, its error component, which is the ratio of the average sample size and the number of populations. Thus, *Rho* = Sb/(Sb+SW). R_{ST}Calc (Goodman, 1997) was used to calculate the pair-wise R_{ST} estimator, *Rho*. Each statistic was tested for significance with 10 000 permutations.

2.5.2.4 Genetic distance and divergence time

The concept of inter-population genetic distance, generated from microsatellite data, is potentially useful in addressing a wide variety of population genetic questions from population structure and isolation by distance (IBD), to coalescence. Genetic distance measures may be derived from statistics that assume either the IAM or the SMM. However, Paetkau *et al.* (1997) suggest that these models are an oversimplification of the mutational dynamics of microsatellite loci and that at the population level, genetic drift is the primary generator of genetic differentiation. The variance associated with

different distance measures is therefore more important than the mutational model assumed.

As allele sizes appear to be highly constrained (Garza *et al.* 1995), genetic distance is predicted to plateau, beyond which point genetic signal will be lost. The possibility of allele sizes being constrained was investigated by sequencing the microsatellites (see Section 2.4.8) of at least two homozygotes (one for each species) to determine the absolute number of repeats each locus contained. All microsatellites were sequenced in both directions and as a further measure, the number of repeats were compared to that of the published (*E. caballus*) clone for comparison. It was found that allele size is a direct reflection of the number of repeated elements as no mutations were detected in the flanking regions of any of the 12 sequenced microsatellites. Where sequences could not be obtained (loci HTG 5, HTG 11 and UCDEQ 505) the number of repeat elements in the published clone was used as the standard. The frequency of each allele class (a total of 13875 alleles) was plotted on a separate histogram for each of the two studied zebra species, allowing the inference of allele size constraints. Allele frequency data for horses was available for 14 loci including (HMB 1 (M. Binns pers. comm.); HTG 3 and HTG 5 (Marklund *et al.* 1994); HTG 7, HTG 9, HTG 11, HTG 14 and HTG 15 (S. Marklund pers. comm.); LEX 20 and LEX 52 (G. Cothran, pers. comm.); VHL 47 (van Haeringen *et al.* 1998); UCDEQ 505 (Eggleston-Stott *et al.* 1997); TKY 273 (T. Tozaki, pers. comm.) and COR 014 (D. Miller, pers. comm.). Although these data were obtained independently from different samples of horses, they were included for a qualitative comparison.

Genetic distance

Most genetic distance measures that assume the IAM will lose linearity less than 20 000 years after the separation of two independently evolving populations (Paetkau *et al.* 1997). This has detrimental consequences for interspecific phylogenetic and coalescence inference from microsatellite data. It may be possible to maintain linearity for a greater length of time by implementing the SMM into a distance statistic (e.g. $(\delta\mu)^2$, Goldstein *et al.* 1995). However, the variance of this measure is high. As many as 30 polymorphic loci (Takezaki and Nei, 1996) may be required before the signal from $(\delta\mu)^2$ becomes biologically meaningful. Valsecci *et al.* (1997) have also found that $(\delta\mu)^2$ may become inflated when sample populations differ greatly in size. As genetic distance between two populations decreases, allele frequencies in the comparison are expected to become identical. The Y-intercept of a plot of genetic distance by

geographic distance should be zero. A further complication is that within-population genetic diversity (H_E) was found to be inversely correlated with genetic distance measures (Paetkau *et al.* 1997; Hedrick, 1999), thus biasing them downward. It is, thus, also necessary to consider this effect when comparing populations with differing levels of genetic diversity.

Although mountain and plains zebra populations in most of southern Africa are still free ranging, those subspecies occurring in South Africa have been reduced to isolated game reserves or extinction primarily due to the early onset of European colonisation and advances in agriculture in that country. Historically, however, both species were continuously distributed across their respective ranges. As there are no obvious allopatric barriers to gene flow in either species' distribution, phenotypic variation across these ranges was probably clinal (see Section 4.1). The genetic structure of these species is therefore predicted to follow that of the isolation by distance (IBD) model (Wright, 1943) where genetic exchanges are restricted to adjacent populations. Genetic distance at neutral loci is therefore predicted to increase with geographic distance (Slatkin, 1993). The suitability of this model may be tested by assessing the correlation between pair-wise genetic and geographic distances.

In the present study, genetic distances are used to analyse the distribution of genetic heterogeneity, thereby elucidating intraspecific genetic structure among populations of each of the southern Africa's two extant zebra species. The distance statistics used were Nei's (1972) standard genetic distance D_S, the arc distance (D_C) of Cavalli-Sforza and Edwards (1967) and $(\delta\mu)^2$ (Goldstein *et al.* 1995). The advantage of all these distance measures, unlike fixation indices, is that they increase linearly with time. D_S was chosen for its low variance and its ability to estimate fine scale population structure in North American bear populations (Paetkau *et al.* 1997). However unlike D_S, D_C is not affected by within-taxon heterozygosity where the distance between two homoallelic taxa that possess alternative alleles exceeds that for two taxa where one or both are heteroallelic but have no alleles in common. D_C also assumes that divergence is generated by either genetic drift or by variable selection (Cavalli-Sforza and Edwards, 1967). The $(\delta\mu)^2$ distance was chosen as the number of loci used in this study, 15, is relatively high compared to that in other studies where the effect of high variance may have masked the genetic signal (Paetkau *et al.* 1997; Valsecci *et al.* 1997). $(\delta\mu)^2$ also incorporates the SMM and is expected to remain linear as long as this model of microsatellite evolution is adhered to.

Matrices of Nei's standard genetic distance (D_S; Nei, 1972), of D_C (Cavalli-Sforza and Edwards, 1967) and of $(\delta\mu)^2$ (Goldstein *et al.* 1995) were calculated by the program Populations 1.2.26 (Langella, 1999). This program also constructed population trees for each distance measure by the neighbour-joining method (Saitou and Nei, 1987), with 1000 bootstrap replicates. Geographic distances were measured as the distance in kilometres between the central points of each sampling location. For mountain zebras, where migration could only occur along areas of rugged terrain, pair-wise geographic distances were measured as the length in kilometres of the most parsimonious path along mountain escarpments between the central points of two sampling locations. A matrix of pair-wise geographic distances was drawn for each species and tested for isolation by distance against the three measures of genetic distance by a Mantel test implemented in GENEPOP 3.1c (Raymond and Rousset, 1995a). Significance of the Spearman rank correlation co-efficient was determined by 10 000 permutations. Genetic distance was graphed relative to geographic distance and regression lines fitted in Microsoft Excel 97.

Divergence times

The conversion of measures of genetic distance to time since two populations diverged from a common stock may be calculated as described by Ciofi and Bruford in 1999. The first method, based on F_{ST}, was found to underestimate divergence times when compared to palaeogeographic data. Instead, a method using the $(\delta\mu)^2$ statistic and incorporating the mutation rate v allows the calculation of divergence time (T) in generations by $T = (\delta\mu)^2/2v$. Being an inverse function of the mutation rate, the estimate of 2.05×10^{-4} mutations per generation was used (Lehmann *et al.* 1998; Rooney *et al.* 1999). The bootstrapping technique, implemented in Microsat 1.5d (Minch *et al.* 1997) with 1000 replicates, was used to obtain 95% confidence limits on divergence time estimates.

2.5.2.5 Gene flow

Gene flow can be measured from microsatellite data as the effective number of migrants per generation (N_em) between two populations. Pair-wise F_{ST} may be used to estimate gene flow from allele frequencies. A flaw in this approach is that F statistics and their derivatives assume an island model (Wright, 1931) where populations receive migrants randomly and at the same rate from an infinite number of equivalent populations. Furthermore, this method assumes the evolution of microsatellite loci by

the IAM and assumes that populations are in mutation-drift equilibrium. The high proportion of homoplasic allele classes not accounted for by the IAM will underestimate pair-wise F_{ST} and consequently overestimate N_em, except in cases where the time scale of interest is very short and differentiation is a product of drift and not mutation (Slatkin, 1995). This method of determining gene flow was therefore not considered in the present study, as the time scale of interest is likely to be higher than 100 generations. Pair-wise R_{ST} (or *Rho*) may be used to measure gene flow (N_em) as N_em = d-1/4d ((1/R_{ST})–1), where d is the number of populations (Ciofi and Bruford, 1999). Although this method also assumes Wright's (1931) island model, it does incorporate the more realistic SMM of microsatellite evolution and is independent of mutation.

Slatkin (1985) showed that the natural logarithm of N_em is approximately inversely proportional to the natural log of the average private allele frequency p(1). More specifically, ln[p(1)] = -0.505 ln(N_em) – 2.440. This assumes both an island model and a two-dimensional stepping-stone model (Hartl and Clark, 1989) where islands exchange migrants only if they are neighbours.

The N_em estimate of gene flow was determined globally and for each population pair from *Rho* values according to Ciofi and Bruford (1999) and by the private allele method (Slatkin, 1985), the last of which is incorporated in GENEPOP 3.1c (Raymond and Rousset, 1995a).

2.5.2.6 Population assignment tests

Tests that assign individual genotypes to populations may be useful in conservation genetics when dealing with specimens of unknown origin. This is especially advantageous when the taxa under study are difficult or expensive to sample. Providing that there is prior genetic knowledge of possible source populations, a log-likelihood method may be implemented to determine population assignment. The log-likelihood of a multilocus genotype is computed, assuming that the individual originates from its designated population. A likelihood score is obtained for each locus from allele frequencies of the designated population, and assuming that each locus is independent, the global likelihood product is then calculated.

Population assignment tests may also be used as an indirect measure of structure between populations. If the multilocus genotype of an individual is correctly assigned to

its designated population, it can be inferred that allele frequencies in this population are different to those in the other populations in the test. At random, the percentage of individual multilocus genotypes that are correctly assigned will be the number of times 100% is divisible by the number of populations. Log-log plots of likelihood scores for two populations have been used to hypothesise dispersal (Waser and Strobeck, 1998). However, Schneider *et al.* (2000) advise that the interpretation of such results in terms of gene flow may be difficult and hazardous. Assignment tests were computed by AGaR$_{ST}$ 2.8 (Harley, 2002) as log-likelihood probabilities of the occurrence of each multilocus genotype in its designated population. Individuals were deemed correctly assigned if they were 95% more likely to be assigned to their designated population than to the next most likely population.

2.5.2.7 Principal component analysis

Principal component analysis is a technique in which microsatellite allelic data are subjected to a multivariate analysis that presents the data in the form of its principal components. This simplifies the data with minimal loss of information. The principal components are those dimensions that best account for all the variance in the data. A goodness-of-fit statistic called the inertia (I) is a measure of how much variance each of the principal components accounts for. Ideally, the majority of the data should be accounted for in the first two principal components. This allows for a qualitative graphical representation of the data in two-dimensional space. The program PCAgen 1.2 (Goudet, 1999) was used to perform principal component analysis on the allelic data. The significance of the inertia of the first two dimensions from zero was assessed by 1000 permutations.

2.5.2.8 Effective population size

The effective population size (N_e) is the size of an idealised population that has the same probability of identity as the actual population being studied (Crow, 1986) or the size of a population that would lose heterozygosity at a rate equal to that of the observed population. N_e is affected by sex ratio and may be obtained directly from demographic or census data. Indirect genetic estimates of N_e allow the assessment of the congruence between population genetic data and real population numbers. Indirect N_e maybe calculated using one of two derivations, each of which assume a different model of microsatellte allelic evolution. Assuming the IAM, Kimura and Crow (1963)

define the effective population size as $N_e = H_E / 4v(1 - H_E)$, where H_E is the expected heterozygosity and v is the mutation rate. Ohta and Kimura (1973) account for mutation and assume the stepwise mutation model thereby deriving $N_e = [(1/(1 - H_E)^2) - 1]/8v$. Although the stepwise mutation model is the more plausible method of allelic generation, both estimates of N_e are reported here for comparison. These were calculated from the data by AGaR$_{ST}$ 2.8 (Harley, 2002), assuming that mutation at microsatellite loci occurs at the rate of 2.05×10^{-4} per generation (Lehmann *et al.* 1998; Rooney *et al.* 1999). Direct N_e values were determined from demographic data by the formula: $1/N_e = 1/4\ (1/N_eF + 1/N_eM)$, where N_eF is the number of breeding females and N_eM is the number of breeding males in a population. The majority of the zebra populations in southern Africa are not intensively managed and do not have associated demographic records. Demographic data on the sex ratio and percentage of breeding adults obtained from two intensive population studies within the southern African sub-region, of mountain zebra at the Mountain Zebra National Park (Dorgeloh, pers. comm.) and of plains zebra at the Kruger National Park (Smuts, 1974), were used to estimate the direct effective population size from population census counts.

2.5.2.9 Population bottlenecking

The precise knowledge of demographic histories of major Cape mountain zebra (*Equus zebra zebra*) populations offers the rare opportunity for directly testing various hypotheses for the detection of recent population genetic bottlenecks in a large mammal. Population bottlenecks may be detected in natural populations using two very different, statistical approaches. The first and more reliable method, developed by Garza and Williamson (2001), relies on the hypothesis that, if an outbred, genetically diverse population undergoes a bottleneck event, it will lose alleles such that the ratio (M) of the number of alleles present to the spread of alleles will decrease. Significant population bottlenecking can be inferred if M falls below 0.81 (Garza and Williamson, 2001). The program AGaR$_{ST}$ 2.8 (Harley, 2002) was also used to compute the M ratio (Garza and Williamson, 2001) for the mountain and plains zebra data.

A second method implies recent population bottlenecking by detecting heterozygote excess (Luikart *et al.* 1998). Both the IAM and the SMM may be used, although there is evidence to suggest that heterozygote excess is more likely under IAM (Comuet and Luikart, 1996). The method is computation-intensive, requiring the distributions of heterozygosities for each sample population and locus, which are then tested for

significant heterozygote excess with a Wilcoxon sign-rank test (Luikart *et al.* 1998). The method requires a large number of polymorphic loci and has been shown to be less sensitive in the detection and estimation of the severity and magnitude of population bottlenecks (Spencer *et al.* 2000). Furthermore, since allele frequencies soon randomise after a population bottleneck, the effect of heterozygote excess does not last as long as that of *M*, where alleles are lost. The method of Luikart *et al.* (1998) is used as a comparison with Garza and Williamson's (2001) test. Bottleneck 1.2.02 (Comuet and Luikart, 1996) was used for testing the distributions of heterozygosities for significant heterozygote excess, using a Wilcoxon sign-rank test.

2.6 Mitochondrial DNA

One to five individuals representing each plains and mountain zebra population were chosen at random and 1.4 kb of the control region of their mitochondrial genomes was amplified. Many samples isolated from dry-salted skin and tanned or weathered samples, while containing amplifiable microsatellite loci (with size range of 67 - 219 bp), failed to deliver the above much larger mitochondrial PCR product owing to DNA degradation. Consequently, some populations such as Gamka, Erongo and Kamanjab, are represented by a single control region sequence.

Amplification primers HorThr (5' CATTACCCTGGTCTTGTAAACC 3') and Hor12S (5' GGCTAGGACCAAGCCTATGTG 3') were used as by Oakenfull *et al.* (2000). PCR was completed in a final reaction volume of 50 μL comprising 1 x polymerase reaction buffer (16 mM $(NH_4)_2SO_4$, 67 mM Tris-HCl (pH 8.8), 0.01% Tween-20; Bioline), 3 mM $MgCl_2$, 0.25 mM DNTPs (AEC Amersham), 0.3 μM end-labelled forward primer, 0.3 μM reverse primer, 0.05 U/μL Taq polymerase (Bioline) and 100 ng template genomic DNA.

Thermocycling was carried out on a 96-well Stratagene Robocycler. The reaction was cycled once at 90 °C for 2 min and 96 °C for 3 min, then 35 times at 96 °C for 15 sec, 50 °C for 30 sec and 72 °C for 60 sec and finally extended once for 120 sec at 72 °C. The reactions were gel purified and sequenced as described in Section 2.4.8 and up to 600 bp of the 1.4 kb control region fragment were obtained in the forward direction using the original amplification primer. The control region of equids contains a minisatellite repeat region near its 3' end. The length of each repeat element was found

to be exactly 24 bp in all mountain zebras and 12 bp in all plains zebras. A 18-mer internal primer, designed to anneal at 545 bp in the forward direction was used to sequence up to 781 bp in mountain zebras and 756 bp in plains zebras at which point the repeat array was encountered. Amplifying through this region in the forward direction proved impossible as the polymerase usually began to malfunction 216 – 228 bp into the minisatellite. Sequencing the complementary strand in the reverse direction was also problematic. An array of G's encountered about 50 bp from the annealing point terminated polymerase function in all cases. Therefore only one strand (5' –3') was used in this study. Sequence chromatograms were checked by eye and reading errors were corrected. Ambiguous sites were treated as missing data. Sequences were aligned using the program DaPSA 4.91 (Harley, 2001). DaPSA 4.91 was also used to format aligned sequence into input files for other programs.

2.7 Mitochondrial DNA analytical techniques and software

2.7.1 Sequence variation and rate heterogeneity

Mitochondrial sequence variation is measured by haplotype diversity (the proportion of distinct haplotypes in any sample; Nei, 1987) and nucleotide diversity (π, the average number of nucleotide differences as a proportion of the entire length of sequence; Nei, 1987). This last-named parameter is affected by differences in substitution rate in different sequences. Rate heterogeneity is assumed to be gamma distributed, with a defining shape parameter (α): the larger the α value, the smaller the rate heterogeneity in the analysed sample. Both distance and maximum likelihood methods are able to estimate nucleotide divergence while incorporating rate heterogeneity, allowing for a more accurate estimation of divergence estimates. Haplotype diversity was calculated by DNAsp 3 (Rozas and Rozas, 1999). Heterogeneity of the gamma distributed nucleotide substitution rate was calculated from each of the two specific data sets by Tree-puzzle 5.0 (Schmidt *et al.* 2000). This enabled corrected estimates of other genetic parameters to be made, such as nucleotide diversity π, nucleotide divergence (D_{xy}, Nei, 1987) as well as their standard errors, calculated by MEGA 2.1 (Kumar *et al.* 2001) and Tree-puzzle 5.0 (Schmidt *et al.* 2000).

2.7.2 Mismatch distribution

Population size changes may be inferred from the distribution of the observed number of differences between pairs of haplotypes which is the mismatch distribution of Rogers and Harpending (1992). When populations are in demographic equilibrium the mismatch distribution is multimodal and ragged but may be unimodal and smooth in populations that have experienced a recent demographic expansion. This provides a convenient qualitative method of assessing recent demographic history. Mismatch distribution analysis was implemented in DNAsp 3 (Rozas and Rozas, 1999).

2.7.3 AMOVA

The analysis of variance framework was originally designed for the analysis of gene frequencies in molecular haplotypes (Excoffier *et al.* 1992). This technique, coined AMOVA, has already been outlined (Section 2.5.2.1) and was carried out in Arlequin 2.000 (Schneider *et al.* 2000) to test for the significance of hierarchical groupings of southern African zebra populations, using 10 000 permutations.

2.7.4 Isolation by distance

Pair-wise Kimura 2-parameter distances (Kimura, 1980) between populations were computed by the program MEGA 2.1 (Kumar *et al.* 2001) and tested against pair-wise geographic distances with a Mantel test as in Section 2.5.2.4.

2.7.5 Phylogenetics

Phylogenetic tree reconstruction provides a concise and graphical representation of the relative amounts of nucleotide divergence in a group of haplotypes, so enabling the inference of evolutionary history. Originally developed to analyse morphological data, most phylogenetic inference employed either numerical phenetic (Sokal and Sneath, 1963) or cladistic (Hennig, 1966) based methods. Phylogenetic analyses of the evolutionary information in the DNA sequences saw the pioneering of maximum likelihood methods (Felsenstein, 1981) for tree reconstruction. The benefit of maximum likelihood (ML) over parsimony, the latter which is able to use only nucleotide sites that have at least two different base residues with both states being represented at least twice, is that it incorporates most of the data available in DNA sequences. The development of the computationally efficient neighbour-joining (NJ) algorithm (Saitou and Nei, 1987) also brought about a resurgence in the use of distance methods in the

construction of phylogenetic trees from molecular data. Although much more computationally intensive, ML tends to outperform NJ when the nucleotide substitution rate is variable (Saitou and Imanishi, 1989). In the present study, mitochondrial control region lineages were used to determine intraspecific population history and structure. The simpler distance based methods of tree reconstruction were neglected here in favour of ML. Maximum likelihood phylogenetic trees with clock-like branch lengths were generated by quartet puzzling (Strimmer and von Haeseler, 1996), estimating rate heterogeneity directly from the data and assuming the HKY substitution model (Hasegawa *et al.* 1985). Significance of groupings was determined from the quartet puzzling support value for each internal branch. Maximum parsimony (MP) of Eck and Dayhoff (1966) was the cladistic method used as it uses discrete characters and can also accommodate insertion/deletion events, treating these as a binary character, and may therefore enhance the detection of an evolutionary signal in variable mitochondrial control region sequences. A maximum parsimony (MP) hypothesis was generated for each species by implementing the heuristic search algorithm in PAUP* 4.0 (Swofford, 1998). Branches of MP trees were tested for significance with 500 bootstrap replicates. In all cases, the horse (*E. caballus*) was used as the outgroup. The 50% majority rule was imposed on all resultant tree hypotheses to obtain consensus trees.

2.7.6 Coalescence

Owing to the uniparental inheritance and non-recombination of the mitochondrial genome, linearity of distance measures is usually not problematic at the intraspecific level. Calibrating particular divergence values with the corresponding fossil record produces a linear molecular clock which enables the dating of pertinent evolutionary phenomena. However, the distance statistic used will depend on the nature of the sequence data. Coalescence of zebra maternal lineages was estimated as in Eizirik *et al.* (2001), using both distance and ML estimates of intraspecific nucleotide diversity, π. Owing to a relatively large fragment length and high substitution rate heterogeneity among control region sequences, the gamma corrected Kimura 2-parameter distance (Kimura, 1980) was used, despite a higher variance for this statistic (Nei and Kumar, 2000). Furthermore, the proposed distance is particularly robust to the high transition/transversion ratios exhibited by the sequence data. Gamma-corrected nucleotide diversity was also estimated according to maximum likelihood as in Oakenfull *et al.* (2000). Two horse (*E. caballus*) sequences were used as the outgroup for both mountain and plains zebra data sets. Average divergence (D_{xy}, Nei, 1987) of

the mountain and plains zebra data sets to the outgroup taxa was also calculated using both distance and ML methods, but corrected by using the α parameter estimated by Oakenfull *et al.* (2000) for the control region sequences across the Equidae. The estimation by Oakenfull *et al.* (2000) of the emergence of an *E. quagga-E. caballus* common ancestor a maximum of 2.3 Mya was used to calibrate both zebra-outgroup divergence estimates, thereby allowing for the estimation of coalescence time for each species.

Chapter 3: Genetic drift and population structuring in the mountain zebra (*Equus zebra*)

3.1 INTRODUCTION

Mountain zebras (*Equus zebra*) were the first zebras described in science by Linnaeus when he introduced his binomial scientific nomenclature in 1758. The species was re-described by Gray (1852) as *Hippotigris*. The "mountaineering hippotigrines" (Woods, 1960) are the smallest of the zebras with a mass of 234 – 298 kg and they stand 124 – 150 cm at the shoulder (Skinner and Smithers, 1990). Their ground colour is white, overlaid with black or chocolate brown stripes, which are anteriorly narrow and numerous but become wide and dark at the hindquarters.

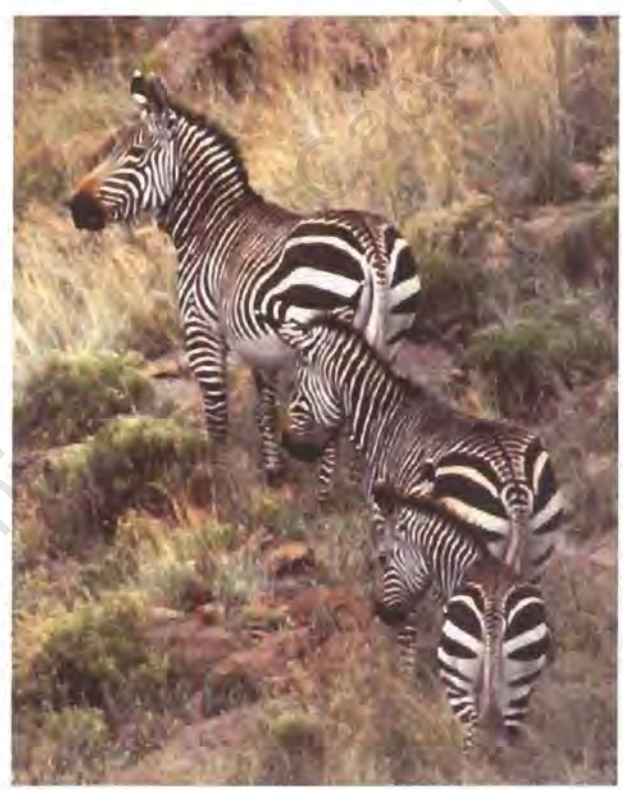

Figure 3.1. Family group of mountain zebras at the Mountain Zebra National Park. Distinguishing features depicted here include the grid-iron pattern on the croup, the dewlap ventral to the throat, an orange-brown muzzle and large, rounded ears. Picture copyright: A. Bannister; *The National Parks of South Africa*, Struik, 1983

3.1.1 Distinguishing features

Mountain zebras possess a distinct grid-iron pattern on the croup which continues posteriorly down the tail (Fig. 3.1). *Equus zebra* may be further distinguished from *E. quagga* in that it has an orange-brown muzzle, larger rounded ears (21 –23 cm long) and a distinctive dewlap on the ventral side of the neck (visible in the highest animal, Fig. 3.1). The mane in *E. zebra* starts almost exactly between its ears whereas that of *E quagga* begins further forward in front of the ears. The hooves of mountain zebras are harder and tend to grow at a much faster rate owing to excessive wear in their type habitat (Skinner and Smithers, 1990). Therefore, mountains zebras are unable to inhabit flat plains for long periods without experiencing overgrowth of the hooves.

3.1.2 Ecology

Like all equids, mountain zebras are bulk grazers. Because they are out-competed on the plains, mountain zebras are restricted to high altitude areas which offer adequate amounts of graze, their preferred graze being the rooigras (*Themeda triandra). Cymbopogon plurinoides*, *Heteropogon contortus*, *Setaria neglecta*, and *Enneapogon scoparius* are also grazed, but to a lesser extent (Grobler, 1983). The species may occur at altitudes as high as 2000 metres (Millar, 1970a) but are usually found on lush slopes and plateaux at somewhat lower elevations. In winter months, when mountain grazing is limited, mountain zebras move to lower altitudes, where they often come into contact with humans. Mountain zebras are completely dependent on water (Joubert, 1971) and often make extensive daily journeys from their grazing grounds down to kloofs and streams. Kloofs as well as caves and ridges offer protection during harsh winters.

Mountain zebras are gregarious and diurnal. Family groups consist of a stallion and up to five mares and foals. Foals of both sexes voluntarily leave the family group at about 22–25 months (males) or during their first oestrus (females) and form mixed-sex "bachelor" groups (Penzhorn, 1975; Lloyd and Rasa, 1989). Although bachelor groups sometimes attach themselves to family groups (Joubert, 1971), mating within a family group by a male other than the herd stallion has never been documented as females within a group are usually highly aggressive to foreign males. Groups are not territorial and large home ranges may overlap (Penzhorn, 1988).

At about five years of age, stallions are able to form herds (Penzhorn, 1984b) and mares produce their first foals at age four to five (Penzhorn, 1985; Penzhorn and Lloyd, 1987) although this may take place as early as three years in some cases (Joubert, 1974). Fertility does not appear to decrease with age and mountain zebras are able to reproduce up to the age of 30 years (Jones, 1993). Generation time is therefore approximated at 16.5 years – the mean of the first and last age of reproduction. Gestation lasts for approximately one year (Joubert, 1974). Foals may nibble grass as early as three days after birth but are fully weaned at 12–14 months. Mortality of foals is high in South African populations studied (Lloyd and Rasa, 1989) but low in Namibian populations (Skinner and Smithers, 1990).

Called "wildepaarden" by the Dutch settlers, mountain zebras were usually seen in association with red hartebeest (*Alcelaphus buselaphus*), ostriches (*Struthio camelus*) and black rhinoceros (*Diceros bicornis*). Their primary predators are the lion (*Panthera leo*) and the spotted hyaena (*Crocuta crocuta*) although it is only the Namibian populations that are now subjected to predation by these species. Leopards (*Panthera pardus*) prey on foals throughout the range of the mountain zebra.

3.1.3 Distribution and early history

Equus zebra once inhabited the continuous chain of escarpment which runs along the west coast of southern Africa, separating the Namib desert from the central plateau; from the Mossamedes district in south-west Angola, south through the rugged Kaokoland, Damaraland, the Khomas Hochland, the Naukluft Mountains and the Huns Mountains, across the Gariep River to the Kamiesberg and the fold mountains of the southern and eastern Cape (see Fig. 3.2).

The San called the fully striped wild horses *Daou* or *Dauw* (Penzhorn, 1988) and depicted them in their rock art. In 1657, five years after European settlers arrived at the Cape of Good Hope, an expedition to the interior encountered the first signs of mountain zebras on the eastern side of Paarl Mountain, in the vicinity of the Great Berg River. The animals were eventually seen in the Twenty-Four Rivers area by members of another expedition in 1658. The "seven horses in the wild state" which were "in good condition", " their colour being dapple-grey" (Woods, 1960). Even today, without the aid of powerful binoculars, the black and white striping of mountain zebras against a backdrop of mountain fynbos, offers considerable camouflage at a distance. In the late 17th century, mountain zebra were common in the mountains of the Western Cape with

documented sightings at Riebeeck's Casteel, Paardeberg, Saronberg, Meerhof's Casteel and the Attaqua Mountains (Bigalke, 1952). As European settlement moved eastward, mountain zebras were found as far east as the Amatola Mountains. They never occurred further east or north into the Drakensberg Range since the altitude of these mountains meant that mountain zebra would have been subjected to frequent snowfalls. From records kept at the Mountain Zebra National Park (Penzhorn, 1984a) and from the Kouga Mountains (Millar, 1970a), mountain zebras, especially the very old and very young, suffer high mortality during particularly heavy snowfalls. Owing to the lower carrying capacity of their type habitat, it is unlikely that they were ever as numerous as *E. quagga*, their plains relatives.

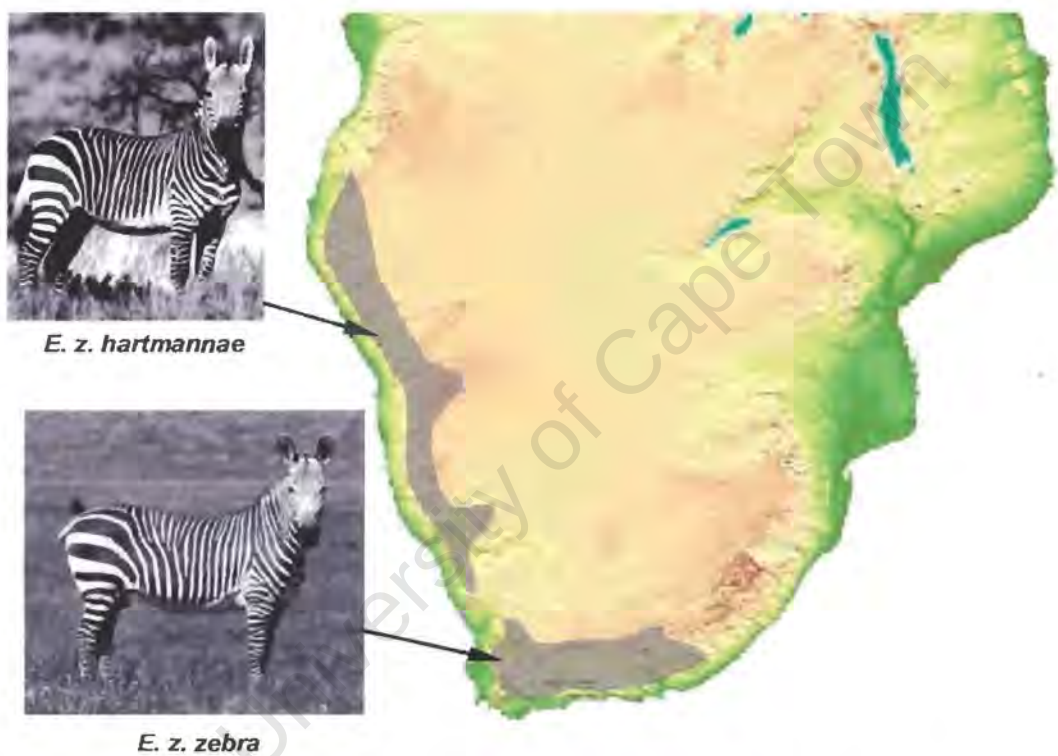

Figure 3.2. Historic distribution of mountain zebra in southern Africa. Hartmann's mountain zebra (*E. z. hartmannae*) ranged along the Namib Escarpment from Mossamedes to the Kamiesberg. Cape mountain zebra *(E. z. zebra*) occurred in the fold mountain region south of the Great Karoo from the Roggeveldberg to the Amatola Mountains.

Governor van der Stel's expedition to Namaqualand in 1685 was the first to describe mountain zebras in the Kamiesberg (Bigalke, 1952). In German South West Africa

(now Namibia), the explorer Sir James Alexander, documented "wild horses" in the Karas Mountains in 1837 (Woods, 1960). On an exploratory trip in 1896, Dr Georg Hartmann shot a zebra between the Unilab and Hoanib Rivers in the Kaokoveld. He sent the skin to his wife, who donated it to the Natural History Museum in Berlin. Two other specimens subsequently arrived from German South West Africa, and after inspection by Professor Paul Matschie, these northern mountain zebras were pronounced as a new subspecies in 1898. He named the mountains zebras occurring in the Kamiesberg and north of the Gariep River *Equus zebra hartmannae* after Frau Hartmann. The mountain zebras that occurred in the fold mountains south of the Karoo and northwards to the Cedarberg and Roggeveldberg became known as Cape mountain zebras (*Equus zebra zebra*). The Knersvlakte, an arid plain north of Vanrhynsdorp approximately 150 km across, is thought to allopatrically divide the two subspecies. Although the subspecific status has never been reappraised since Matschie (1898), the delineation has recently been justified on the basis of the following characteristics (Novellie *et al.* 2002): *Equus zebra hartmannae* is thought to be larger than *E. z. zebra*, black stripes are wider in *E. z. zebra* and the mane extends further forward between the ears in *E. z. hartmannae*. Rau (2002) suggests that these phenotypic differences in mountain zebras are plastic and that the extant Cape mountain zebra phenotype is a product of either inbreeding or local environmental conditions.

3.1.4 Hartmann's mountain zebra

Very little is documented about the northern subspecies as appreciable settlement in German South-West Africa was limited to just prior to the turn of the 20th century. Due to very arid and difficult conditions, human population density in the mountainous areas of the Western Escarpment was low. The Namibian Ministry of Environment and Tourism (MET) has conducted censuses to determine mountain zebra numbers since 1972. Although human habitation and pastoralism have increased, especially in the Central region, mountain zebra numbers in this area appear not to have been affected (Novellie *et al.* 2002). Further north in Damaraland and Kaokoland, mountain zebras are essentially free ranging to the Kunene River and to Angola beyond. In the vast area south of the Naukluft Mountains, *E. zebra* numbers have declined (Novellie *et al.* 2002), although they are still found as far south as the Huns Mountains, the Fish River Canyon and across the Gariep River to a small population in the Richtersveld National Park in South Africa. Early explorers documented the Hartmann's mountain zebras of

these southern populations and noted a higher historic population density than is present today. The reduction in numbers of southern Hartmann's mountain zebra was probably a result of human population radiation from either the original settlement at the Cape or German colonisation from the southern port of Luderitz, or both. Despite an "endangered" listing for mountain zebras in the IUCN's Red Data Book of threatened species (Hilton-Taylor, 2000), poaching continues to have a detrimental effect on these small mountain zebra populations.

With the advent of cattle and sheep farming on the Namibian Plateau, fences were erected. Such fenced pastoralism is most concentrated in the Central region of Namibia, which encompasses the prime mountain zebra habitats of the Khomas Hochland, the Gamsberg and the Erongo Mountains. The MET suspects that fences significantly hinder the natural migration routes of these central Namibian mountain zebra populations. Adjacent fenced-in populations in the Central region are expected to be more differentiated from each other than adjacent populations in other regions. Furthermore, genetic structuring between mountain zebras in the Northern, Central and Southern regions may be apparent as all gene flow between the south and north (and *vice versa*) must pass through the Central region.

3.1.5 Cape mountain zebra

The Cape government prohibited the hunting of mountain zebras in 1742 (van der Merwe, 1962) but the size of the Cape Colony made enforcing this protection impossible. Over a hundred years later, in the days of British governance, the enactment of similar measures saved the bontebok (*Damaliscus pygargus*) from extinction.

In the belief that Cape mountain zebras competed with livestock for grazing and broke fences which divided their home ranges, they were hunted relentlessly (Penzhorn, 1975). To farmers in the southern Cape, mountain zebras amounted to little more than vermin and were dealt with accordingly. They were also hunted for their valuable hides; Bryden (1889) describes instances where farmers would succeed in driving zebras off precipices and would collect the skins at leisure. At the same time, there was also a growing trend to capture foals with the aim of breaking them to harness. A premium of £20 was offered for foals delivered to Cape Town. "A good many" were exported from the Cape to the French Colony of Mauritius (Bryden, 1889). With these almost

insurmountable pressures, it is a wonder that the subspecies survived into the 20th century.

By 1900, two centuries of hunting and exploitation had impacted greatly on the mountain zebras that competed with farmers in the southern Cape. *Equus zebra zebra* numbers had been decimated to such an extent that they were eradicated in most of their former range. Information on the exact whereabouts of surviving populations was vague, with zebras reported to have been seen existing in small troops "here and there" (Woods, 1960). The Cape government listed mountain zebras as Royal Game and a permit was required to hunt or capture one in the Cape Province (Butler, 1999). However, the animals were already so rare that the few farmers with mountain zebras on their land could afford to tolerate them. The children of some of these farmers grew up to take a keen interest in the preservation of their own Cape mountain zebras.

In the years after the Boer War, a circus run by the American cowboy, Texas Jack, toured South Africa. Butler (1999) provides a detailed account of the affairs of Texas Jack while his circus passed through Cradock in 1903. There is much evidence that one of his employees, Will Rogers, had 'roped' mountain zebras "for amusement" during this period (Butler 1999). A picture of the entire circus crew, taken in Port Elizabeth in May 1903, less than a month after it toured through Cradock, features prominently a single Cape mountain zebra in a set of harnesses, with a baboon on its back. The cowboys chased the zebras on horseback until they became exhausted. The fatigued zebras would then offer no resistance to the cowboys who lassoed and even rode them.

At this time the largest and seemingly undisturbed population of *E. z. zebra* existed in that part of the Outeniqua Mountains between the Montagu Pass in the east and the Saffraan River in the west. Ockert Heyns owned the farm Groot Doringrivier, which encompassed much of the land on the northern slopes of the central Outeniquas in the Beveraaskloof area. Cape mountain zebras proliferated on the rugged upper reaches of Heyns's land where he sometimes grazed his cattle. In an attempt to tame his zebras another American cowboy, named Rattery, was hired to capture some of the zebras on Groot Doringrivier. Using the lasso technique, Rattery succeeded in capturing seven, but failed to break them into harness. The captured zebras remained in a kraal on Groot Doringrivier for several weeks, until they died. Heyns attracted the attention of Franz Katzenstein, a wealthy German businessman. In their first

transaction dated May 1909, Katzenstein purchased from Heyns two Cape mountain zebra (see Fig. 3.3), for the combined sum of £300 (letter from F. Katzenstein to O. Heyns, supplied by P. Heyns). Katzenstein visited the Heyns family on at least five subsequent occasions between 1909 and 1926, purchasing "about a dozen zebras" each time (I. Groenewalt, pers. comm.). Millar (1970b) reported that a further 30 animals were captured from this part of the Outeniqua Mountains in 1928 and sent to Europe. Most did not survive the journey. Heyns's son, Paul, finally put a stop to the exportation and dedicated his efforts to the preservation of the Beveraaskloof zebras. However, numbers continued to decline as hunting and farming in the area became more intensive (van der Westhuizen, pers. comm.). From an estimated 150 animals in 1930, only 40 Outeniqua zebras existed in small herds on the northern slopes by 1937 (Millar, 1970b). At the western end of the Outeniqua range, the small herd that existed on the farm Saffraanrivier died out by about 1961.

Figure 3.3. Franz Katzenstein on his first visit to Groot Doringrivier in May 1909. Also pictured are Walter and Dora, the two Outeniqua Cape mountain zebras purchased from Ockert Heyns. Only the male on the left is sternally striped. Picture supplied by P. Heyns, Mount Hope, Beveraaskloof.

In other areas, the government-subsidised fencing scheme, initiated in the 1920s, radically altered the ability of mountain zebras to roam freely the areas they had in the past (Butler, 1999). In the few areas where mountain zebras were tolerated, their

routine fence breaking quickly turned many farmers against them and permits to hunt them were handed out liberally. More than 100 Cape mountain zebra were shot in the Cradock District alone (Boroughs, 2000). By 1926, Cape mountain zebra numbers were further reduced to "about a hundred in various places" (Lang, 1935; cited in Woods, 1960). Besides those animals in the northern Outeniqua Mountains, the existence of small herds was confirmed in the Kouga Mountains, the Kammanassie Mountains and the Outshoorn, Cradock and Sutherland districts (Woods, 1960).

A population of Cape mountain zebras existed in the Sneeukrantz section of the Roggeveld Escarpment, 40 km west-northwest of Sutherland. These animals were isolated for many years and for some time managed to remain in good numbers. The drought of the early 1930s took a heavy toll as weakened zebras were easily caught and eaten. The last surviving stallion was shot in about 1935 as he was deemed "rather rough" with the donkeys and horses with which he was forced to associate (Woods, 1960).

It was now obvious that unless drastic measures were taken, the wildepaarden that had roamed the mountains of the southern Cape would follow the quagga (*Equus quagga quagga*) into extinction. The government was requested twice to declare some of the large tracts of State Land in the Outeniqua Mountains between the Robinson and Montagu Passes a reserve to protect what was still then the largest and best preserved population of Cape mountain zebras (Lloyd, 1984). Finally in 1936, the Minister of Lands, General Jan Kemp declared, "No, they're just a lot of donkeys in football jerseys" (Pringle, 1982), thus sealing the fate of the Outeniqua mountain zebras.

3.1.6 The protected population

It is fortunate that farmers like Paul Heyns took an interest in the preservation of the mountain zebras on their land. In the Bankberg, the Michaus brothers of the farm Doornhoek and the Lombards of Waterval emulated these efforts. These farmers protected healthy numbers of mountain zebras on their farms as had their fathers before them. There were other smaller farms in the Bankberg where mountain zebras were also known to occur, albeit in smaller numbers. In 1937, two years after their initial response, the government purchased the small 1712 ha farm Babylon's Toren approximately 27 km south west of the town of Cradock and proclaimed it the Mountain Zebra National Park (MZNP). South Africa's fifth National Park was faced with an

immediate crisis in that it possessed six zebras, only one of which was female (Pringle, 1982). In 1938, Michaus and Lombard were approached in an effort to remedy the situation. Both families refused to release their zebras and it was assumed that they were simply reluctant to part with the animals that had come to be their pets. However, the reason for their refusal soon became apparent. The quality of the grazing on Babylons Toren was inadequate to sustain even a small zebra population. The Park possessed very little of the rooigras (*Themeda triandra*) that the zebra found most palatable. Consequently, the mare and the only filly she produced had died by 1946 and by 1950 the population had dwindled to two stallions. Then, in that year, the last mountain zebras in the MZNP were shot and the remains sent to the director of the Transvaal Museum. The Mountain Zebra National Park then existed without any zebras until November 1950, when Lombard donated the animals on Waterval to the nation. A fence parallel to one that spanned the farm Pretorius Kraal was constructed, creating a corridor between Waterval and the MZNP and Lombard's five stallions and six mares were ushered into the park. Within 14 years the population had increased to 25 animals (Penzhorn, 1975). In 1964 the Government purchased Rooiplaat, a lush plateau of rooigras, from the Michaus brothers as well as four other neighbouring farms (Lloyd, 1984) and the Michaus reluctantly parted with their 30 Cape mountain zebras which, like the Lombards, they had preserved for generations. The population at MZNP was then 55 and the largest in existence. In the 6536 hectares of the MZNP they flourished to over 200 in 1979. Since then the population size has been controlled to between 200 and 300 animals by the South African National Parks (SANP) and 'excess' animals have been used to seed other reserves (see Section 3.1.8).

3.1.7 The unprotected populations

Between the 1930s and the 1960s, the other Cape mountain zebra populations that remained attracted little attention but in the winter months when harsh weather conditions forced them to move to the lower slopes and fenced camps, regular sightings were reported from the George-Oudshoorn-Willowmore districts. It is interesting to note that some of these populations, such as those in the Kammanassie and Outeniqua Mountains, inhabited state land under the control of the Department of Forestry but it was the Cape Department of Nature Conservation (CNC) assisted by the South African Mountain Club that initiated an intensive foot and aerial survey of this area at various times from 1967 to 1969.

The period between 1969 and 1972 was the most disastrous in the history of the conservation of the Cape mountain zebra, which at the time was one of the most critically endangered large mammals in South Africa. There were no more than 39 zebras in four populations outside the MZNP, probably representing the most diverse component of the Cape mountain zebra gene pool.

Millar (1970a, 1970b) summarised the results of the CNC census and suggested that all Cape mountain zebras in unprotected areas be relocated to the Gamka Mountains where a reserve was to be created for them. As land in the Gamka Mountains had not been purchased, the unprotected zebras would be relocated provisionally to DeHoop Nature Reserve (*Anon.* 1970/71).

3.1.7.1 Gamka Mountains

The Gamka Mountains are equidistant from Calitzdorp in the west and Oudtshoorn in the east. The range is bordered by the Gamka River which separates it from the Rooiberg in the west, the Saffraan and Moeras Rivers in the east and a valley of approximately 7km in width distinguishes them from the western Outeniqua Mountains in the south. Over 30 Cape mountain zebra were known to exist in the 21 000 hectares of privately owned Little Karoo renosterveld in 1937 but the 1969 aerial survey discovered only 13 in four groups (Millar, 1970b). A farmer owning part of this land with knowledge of the census of the previous year and of the importance of the remnant populations was alleged to have shot seven Gamka zebras in a day (Lloyd, pers. comm.). Within a year of the census, this population was reduced to six animals.

3.1.7.2 Outeniqua Mountains

In 1937, when the MZNP was proclaimed, there were estimated to be over 40 zebras roaming the central part of north-facing slopes of the Outeniqua Mountains. Owing to a rain shadow, precipitation on this leeward side of the Outeniqua Range is typically low. Like the nearby Gamka Mountains, Little Karoo renosterveld is prevalent, with an abundance of *Themeda triandra*, making it ideal for mountain zebra habitation. It is unlikely that mountain zebras ranged onto the wetter windward forested southern slopes which were devoid of grazing. As a result of the government's decision not to protect these animals, hunting continued unabated until a 1967 foot census revealed that just three animals remained. These were restricted to the Beveraaskloof. It is

unlikely that these animals would have survived here were it not for the Heyns family and their intimate association with them.

The last remaining Outeniqua zebras all possessed sternal striping across what is a normally white chest area. By chance, some of the remnant Gamka and Kouga mountain zebras also possessed sternal striping leading one conservationist (Lloyd, 1984) to hypothesise that the Gamka, Outeniqua and the Kouga populations existed in relative isolation from all others and that the sternally striped phenotype could therefore be regarded as an endemic sub-form. Although mountain zebras undoubtedly migrated freely between the three mountain ranges and may even be considered as part of the same population, the majority of the zebras in the Outeniqua, Gamka and Kouga Mountains could not have been sternally striped (Fig. 3.3, page 42 and unpubl. photographs; P. Heyns, pers. comm.). Corroborating this view is the relatively common incidence of sternal striping in the Namibian Hartmann's mountain zebra, suggesting that the phenomenon probably occurred randomly to the same extent in historic Cape mountain zebra populations. The present preponderance of sternal striping in Gamka is more likely a result of recent genetic drift on a phenotypic character than of a more ancient phenotype associated with a distinct evolutionary trajectory.

One of the last three Outeniqua mountain zebras, a sternally striped mare later named "Witch", was caught by the CNC Game Capture Team in November 1970 (Fig. 3.4). Van der Westhuizen (pers. comm.) recalls a helicopter disappearing into the mountains twice but emerging only once with a single zebra on their especially constructed stretcher. Two animals were seen by the capture team but only one was captured by immobilization. All that was mentioned of the other zebra was that "an accident" had occurred. This was not reflected in the official CNC capture report. The mare was relocated to DeHoop Nature Reserve, where she died a year later, foaling only a donkey hybrid. Mountain zebras no longer exist in the Outeniqua Range.

Figure 3.4. The capture of "Witch", the last of the Outeniqua mountain zebras, 1970. The Outeniqua Mountains supported the largest population of Cape mountain zebras 20 years earlier. Picture: van der Westhuizen, Klipdrift, Beveraaskloof.

3.1.7.3 Kouga Mountains

One of the few things known about the Kouga Mountains was that they were home to a few relict Cape mountain zebra (Woods, 1962). The Kouga Mountains are actually three parallel ridges running from west to east, between Willowmore and Patensie in the Eastern Cape Province. The narrow valley of the Upper Gamtoos River separates it from the Baviaanskloof Mountains in the north and the Kouga River Valley separates it from the Tsitsikamma Mountains in the south. In 1937, there were an estimated 25 Cape mountain zebras in the central Kouga Mountains, near the settlement of Studtis (Millar, 1970b). Woods (1962) visited the area and was led to within 50 yards of a male and female zebra. The 1968-1969 census discovered that seven animals remained, all in the vicinity of Studtis. The larger herd consisted of five zebras. The male of the pair seen by Woods in 1962 died in the heavy snowfall of June 1968 but fresh signs of the old female were found.

The CNC capture team was dispatched to the central Kouga Mountains in December 1971. Capture nets were erected in Moordenaarskloof where a breeding pair were

caught and at Kasey's Kloof where a further two males were captured. The nature of the terrain was such that the zebras needed to be immobilized and airlifted on a stretcher. A remaining stallion was located and darted above Kotze's Farm but died soon after the administration of the antidote in a holding pen. The Moordenaarskloof pair were released on DeHoop and the Kasey's Kloof males were taken to the MZNP. None of the four captured Kouga mountain zebras reproduced. The last known record of zebra occurrence in the Kouga Range was in 1982 (Lloyd, 1984) and the population is now believed to be extinct.

3.1.7.4 Kammanassie Mountains

The Kammanassie Mountains, an inselberg standing 4000 feet above sea level, lies between the towns of Dysselsdorp in the west and Uniondale in the east. The range is steep with deep ravines and kloofs. The upper sections of the Kammanassie Range were state owned, very rugged and under the jurisdiction of the Department of Forestry. It was discovered that 16 zebras divided into four separate herds with two lone stallions inhabited these mountains just above the farms Klein Geluk, Welgevonden and Leeublad (Millar, 1970a). These animals were exposed to the wrath of owners of these farms each spring when they moved down to graze on their green wheat fields. It was suggested that the chief reason for the decline in Kammanassie zebra numbers, from an estimated 30 animals in 1956 (Skead, 1956), was their killing by farmers (Millar, 1970a). That they were able to survive in these mountains at all is a consequence of the ruggedness and inaccessibility of the terrain. To curb this rate of loss, the relocation of Kammanassie mountain zebras became a priority.

As a result, in November 1970, the CNC sent a game capture team to the Kammanassie Mountains. The team succeeded in driving two male zebras into nets on the farm Leeublad. The stallions, named "Tom" and "Jerry", were transported to a fenced 400 ha camp on DeHoop Nature Reserve. Over a year later, Jerry died of unknown causes and for no apparent reason, Tom was moved to the MZNP. The population at MZNP at the time was over 100 animals. Tom, a large robust animal, succeeded in becoming a herd stallion, siring up to ten foals (Novellie, *in litt.*). In September the following year, a second trip to the Kammanassie Mountains was undertaken by the same game capture team. On this trip, the new drug cocktail of M99-hyascine-acetyl promazine, initially used in the Outeniqua and Kouga captures was utilised. Nine zebras were captured but three died as a consequence of improper dosage and of stress. A further two were "lost" and one stallion was killed when the

wire cable connecting the transporting stretcher to the helicopter snapped at 600 feet. Only two of the nine animals survived and were successfully taken to DeHoop. The population on the Kammanassie Mountains thus also dwindled to no more than five animals.

3.1.8 The present situation

The decimated Cape mountain zebra metapopulation persists today as part of three small relict stocks: Cradock (MZNP), Kammanassie and Gamka. Each population suffered gross reduction and apart from the Kammanassie male "Tom's" success at Cradock, the three stocks have remained isolated from each other. It has been less than two generations (30 years) since the Kammanassie and Gamka populations were at their lowest level. However, as all three populations were significantly reduced in number for over 100 years, genetic drift may have affected these small, isolated stocks over a longer period. The duration of these bottlenecks is expected to be reflected in population genetic analyses since prolonged low population numbers and hence high genetic drift will randomly drive alleles to either high frequency and fixation or to low frequency and extinction.

The population at MZNP increased to numbers large enough to make possible the founding of satellite populations. The first of these was at DeHoop Nature Reserve in the southern Cape where seven animals were reintroduced between 1963 and 1975. Large-scale translocation occurred shortly thereafter. In 1978 the first 16 of 76 Cape mountain zebras were moved to the Karoo National Park (KNP) at Beaufort West. These were the first of many translocation events resulting in the redistribution of Cape mountain zebras to no less than six National Parks, 10 provincial reserves and 17 private reserves (Novellie *et al.* 2002). Of the 1600 Cape mountain zebras in existence today, an estimated 1499 (94%) have their origins solely at the Mountain Zebra National Park. This extended network of satellite populations will be referred to collectively as "Cradock-derived" to distinguish them from the original population at the MZNP.

Given that the MZNP population suffered a genetic bottleneck of 19 animals, satellite translocations, while expanding the geographic distribution of Cape mountain zebras, were short-sighted in that the genetic complement of each population was unknown. Each translocation can therefore be seen as a secondary bottlenecking event, further

exacerbating the effect of genetic drift and resulting in a number of smaller, more inbred populations of Cradock-derived mountain zebra. These populations were intermittently restocked by more MZNP animals and latterly from the increasing KNP population.

The 49 430 ha Kammanassie Nature Reserve was proclaimed in 1970 to protect the last remaining mountain zebras of that area. Since then the founding population has grown slowly to its largest post-bottleneck size of over 40 in 1999 (G. Cleaver pers comm.). A major setback to the survival of this stock has recently occurred. The Department of Water Affairs began abstracting artesian water from the Kammanassie Mountains in 1993 for the nearby town of Dysselsdorp. Since then, almost all of the zebras' natural watering points have dried up, resulting in them having to undertake hazardous treks down to farmland in the south of the mountain range. In the last three years, four animals (a mare and three foals) have been killed in a variety of mishaps that occurred during this journey. The population therefore now stands at 38 individuals. When only breeding individuals are counted, a much smaller effective population size results (Table 3.1). The rugged and inhospitable nature of the Kammanassie Range poses another problem in that the zebras in the east have no contact with those in the west. Proactive management is required to make sure that all individuals contribute to the gene pool.

Land totalling 9428 ha in the Gamka Mountains was finally purchased and the Gamka Mountain Nature Reserve (GMNR) proclaimed in 1971. From its bottleneck of six, the population now comprises 28 animals (T. Barry, pers. comm.) and is the smallest of the three remaining Cape mountain zebra stocks (Table 3.1).

Table 3.1. Social organisation and effective population size in four key Cape mountain zebra populations. Values are given for (Br)eeding (M)ales and (F)emales, subadults and (Bach)elor males. The effective population size (N_e) was determined by the formula $1/N_e=1/4$ (1/Br M +1/Br F). a, T. Barry (pers. comm.); b, G. Cleaver (pers. comm.); c, W. Dorgeloh (pers. comm.); d, P. Lloyd (pers. comm.)

	Br M	Br F	Subadult M	Subadult F	Bach M	TOTAL	N_e
Gamka [a]	4	16	2	3	3	**28**	12.8
Kammanassie [b]	12	18	2	5	1	**38**	28.8
MZNP [c]	63	107	20	22	6	**218**	158.6
DeHoop [d]	9	19	10	12	20	**70**	24.4

DeHoop Nature Reserve has the only known population derived from an equal proportion of two relict stocks, Kammanassie and Cradock. Mountain zebra population growth at DeHoop has been steady since its founding (Novellie *et al.* 2002). Examining the genetic complement of this population relative to the others may provide clues that could facilitate the efficient management of the subspecies. Since the ill-fated game capture operations of the 1970s, the CNC stopped all movement of animals from Kammanassie, Gamka and DeHoop on the grounds that these populations were too small and sensitive and in order to protect their "genetic integrity". It is thanks to this conservative approach that some of these populations still exist, and both Kammanassie and Gamka have shown positive population growth over the last 15 years (Fig. 3.5). Furthermore non-reproducing bachelor groups have recently formed on these reserves (Table 3.1). However, depending on the genetic status of extant Cape mountain zebra stocks, this 'hands off' approach may need to be revisited.

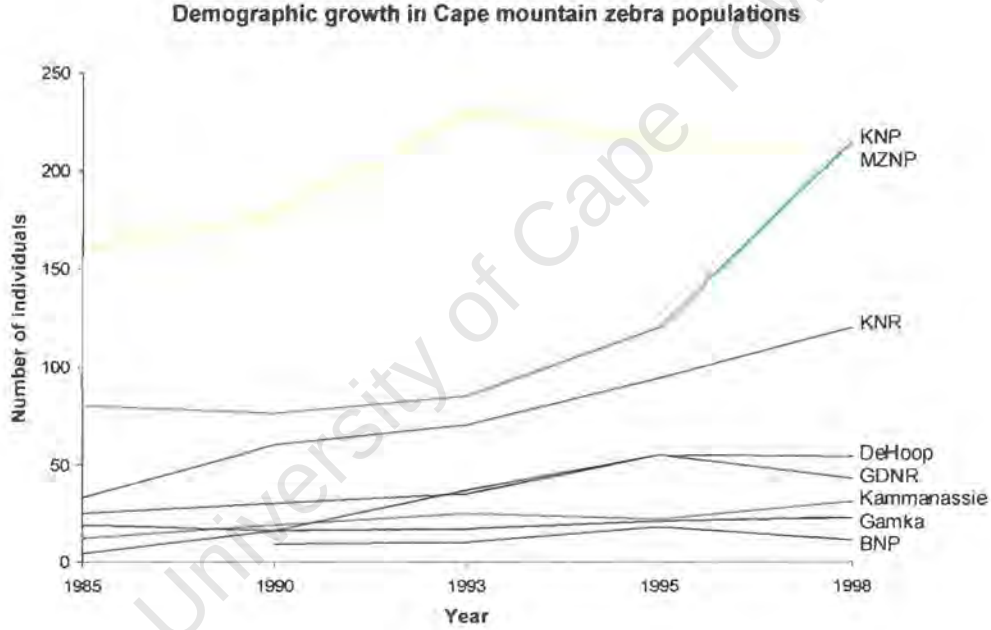

Figure 3.5. Growth over time in studied Cape mountain zebra populations according to data compiled by Novellie *et al.* (2002). KNP, Karoo National Park; MZNP, Mountain Zebra National Park; KNR, Karoo Nature Reserve; GDNR, Gariep Dam Nature Reserve; BNP, Bontebok National Park. Demographic growth graphed only to 1988 as more recent information was not available for all populations. Figures were modified from original data to account for periodic reintroductions to satellite populations. The authors recognise that the abnormally high population growth of KNP for the period 1995-1998 is due to a census error occurring in 1995.

3.1.9 Presence of equine sarcoid

Equine sarcoid is a dermal fibroblastic wart-like lesion which shows variable phenotypic manifestations (Goodrich *et al.* 1998). The suppression of tumour development in fibroblastic tissue is critically dependent on the p53 tumour suppressor protein. The frequency of oncogenic p53 mutations was found to be low in equine sarcoids (Bucher *et al.* 1996) and functional inactivation of p53 by the E6 protein of the bovine papillomavirus (BPV) was suggested as the causative agent. Nasir and Reid (1999) demonstrated BPV gene expression in equine sarcoids thereby providing the first evidence for the direct involvement of BPV in the pathogenesis of sarcoids. Additionally, associations between genes in or near the equine major histocompatibility complex (MHC) and susceptibility to sarcoids has been demonstrated (Lazary *et al.* 1985; Meridith *et al.* 1986). Genetic predisposition for sarcoids has also been suggested by Brostrom *et al.* (1988) and Lazary *et al.* (1994) who describe the association between an autosomal dominant gene linked to equine leucocyte antigens (ELA) and the oncogenesis of sarcoids. All evidence therefore points to a multifactorial aetiology of equine sarcoid.

Although sarcoids do not fit all criteria for malignancy, they are often invasive and recurrent (Martens and de Moor, 1996). The anatomical location of sarcoid lesions may impact on individual fitness. Sarcoids most commonly affect the ventral body regions and the head, neck and the sites of thin skin. Sarcoids on limbs or eyes may inhibit movement and foraging ability and lesions on the genitalia or mammae may affect reproductive ability. Equine sarcoid has been detected in three Cape mountain zebra populations: a stallion in the Gamka population was recorded as having a large tumorous growth (Barry, pers. comm.) and some animals at Bontebok National Park (BNP) and Gariep Dam Nature Reserve (GDNR) exhibit sarcoid neoplasms. The populations at BNP and GDNR were both seeded with 24 Cradock-derived animals from 1986 to 1997 and it is conceivable that at least some of those genetic factors that predispose animals to sarcoid oncogenesis may have been concentrated here as a result of inbreeding.

3.1.10 The role of molecular genetics

Although it is quite clear that the three Cape mountain zebra stocks have undergone extreme population bottlenecking, the status of the resultant gene pools is not known. Such severe population bottlenecks have serious genetic consequences since they drastically reduce genetic variation (Hedrick, 1999) and leave populations open to the effects of inbreeding with continuing loss of genetic diversity due to genetic drift. If left unchecked, inbreeding may reach levels where inbreeding depression (see Section 1.2.1) compromises fitness. The present genetic status of the three key Cape mountain zebra populations as well as that of the entire metapopulation will determine their vulnerability to extinction.

Microsatellite loci are extremely useful in the detection of inbreeding or the loss of genetic diversity as outbred populations are usually highly polymorphic (Bruford and Wayne, 1993). To appreciate the extent to which genetic variation was lost from extant populations of Cape mountain zebra, an estimate must be obtained of the amount of genetic variation present in the metapopulation before genetic bottlenecking. One way of estimating this is by comparing genetic variation in Cape mountain zebra with its closest relative, the Hartmann's mountain zebra (*E. z. hartmannae*). This will help to provide information on the historic genetic diversity of the Cape mountain zebra preceding the population bottleneck.

Although very similar in morphology and distribution, the exact genetic relationship between the two *E. zebra* subspecies has never been examined. Oakenfull *et al.* (2000), in their survey of the control region and 12S rRNA gene of equid mtDNA, noted that *E. zebra* was the most divergent group of stenoid equids, dating the split between *E. zebra* and the *E. quagga*/*E. grevyi* group at 0.67 Mya. Only the subspecies *hartmannae* was sampled by Oakenfull *et al.* (2000) and it was found that haplotypes were shared at the intraspecific level between subspecies of Asian wild asses and plains zebras. Whether or not the same phenomenon occurs within *E. zebra* may be assessed by sequencing the variable control region of the mitochondrial genome. It seems clear that resolving the phylogenetic status of the mountain zebra subspecies is a key component of the present study because it may have consequences for the management of the species, thereby justifying the comparative approach mentioned above.

3.1.11 Aims

The aims of the mountain zebra component of this study are:

a) To carry out a definitive molecular genetic study of Cape mountain zebra (*Equus zebra zebra*) populations and test the hypotheses that populations are inbred and that extreme population bottlenecks have reduced genetic variation.

b) To determine the level of historic genetic variation in Cape mountain zebras by quantifying levels of genetic variability in populations of the closely related Hartmann's mountain zebra (*Equus zebra hartmannae*).

c) To test that hypothesis that there is gene flow occurring between Hartmann's mountain zebra populations in the Northern and Central regions of Namibia.

d) To determine the amount of nuclear and mitochondrial genetic differentiation between the Cape and Hartmann's mountain zebra: testing the hypothesis that there are real genetic differences to substantiate their current subspecific status?

e) To combine all the above genetic information with existing ecological, social and demographic data in order to provide a framework for maximizing and maintaining genetic variation, so ensuring efficient and effective conservation management and the survival of the Cape mountain zebra.

3.2 METHODS AND MATERIALS

Details of the DNA and analytical methods used here are described in Chapter 2.

3.3 RESULTS

3.3.1 DNA isolation and amplification

DNA samples isolated from whole blood or dry salted skins performed well when subjected to PCR amplification. In a few cases, where skins were particularly old, microsatellite amplification products were faint, requiring a prolonged period of exposure.

3.3.1.1 Weathered, faecal and tanned samples

In more extreme cases, modifications (see Section 2.3) to the standard DNA extraction procedure (Sambrook *et al.* 1989) resulted in the isolation of very small amounts (5 – 100 ng) of low quality DNA (A260/A280 < 1.6). This standard SDS-Proteinase K/Phenol-Chloroform method failed to produce tenable DNA isolates from faeces. Three other methods, adsorption by potato flour (Deuter *et al.* 1995), the silica adsorption method (Taberlet *et al.* 1997) and the Qaiamp DNA extraction kit (Qaigen), were tested. Of these, silica adsorption was found to be the method of choice for faecal samples. It was found that sun or oven (70° C) dried faeces produced more amplifiable DNA isolates than any other method of faecal preservation. The extent of nucleic acid degradation appears to decrease when faeces are dried relatively rapidly. However, in the harsh mountainous terrain of the Gamka and Kammanassie Mountains, it was impossible to ascertain how long faecal samples were exposed to the elements or how quickly they dried. Furthermore, it was impossible to determine whether we had sampled the same animal more than once. Almost all of the faecal samples collected from these populations proved to be less than ideal.

Despite the various DNA extraction procedures outlined in Section 2.3, proteinacious contamination was still excessive in tanned skins, dried tissue and faeces. This resulted in poor PCR amplification. In many samples, inconsistent amplification products were obtained; in some cases (weathered: Gamka 1, 2, 6, 7, 8; faecal: Gamka 12 - 19 and Kammanassie 10 - 12; tanned: Gamka 3, 20 and MZNP 1, 2, 3, 4, 5) DNA was extracted from the same material up to five times and each of these extracts amplified no fewer than five times. Individuals were only positively scored

when the same result was obtained in more than 90% of the amplifications. Consequently, only nine (Gamka 2, 4, 5, 6, 9, 10, 11; Kamanassie 1 and MZNP 14) of 12 weathered samples, one (Gamka 12) of 11 faecal samples and four (Gamka 20, MZNP 2, 3, 4) of the seven tanned samples were usable. Figure 3.6 depicts the PCR products obtained from some of these samples.

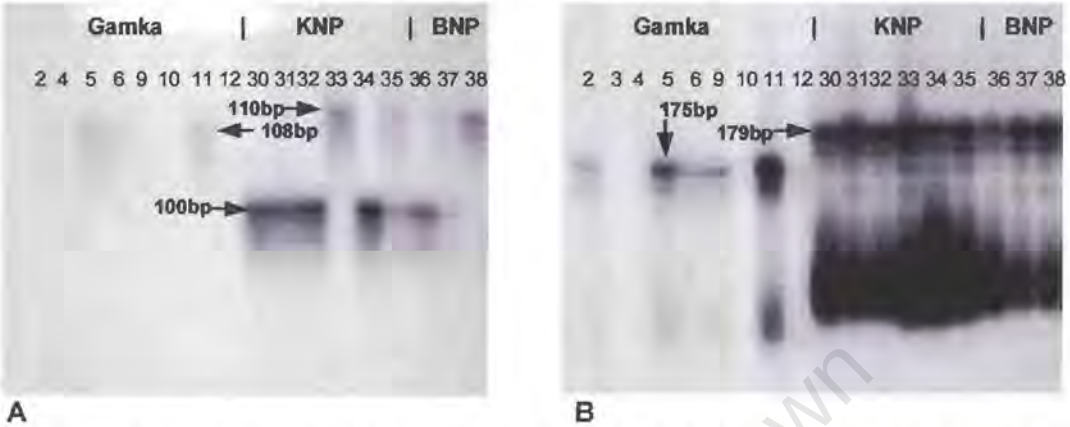

Figure 3.6. Visual comparison of amplification products of both weathered (Gamka) and recent (KNP and BNP) Cape mountain zebra samples at two microsatellite loci. Figure 3.6A, locus HMB1 (Binns *et al.* 1995); Figure 3.6B, locus UCDEQ505 (Eggleston-Stott *et al.* 1997).

Even from these raw data it is apparent that Cape mountain zebra populations have lost genetic variation. Locus HTG 5 (Fig. 3.7A) and locus LEX 20 (Fig. 3.7B) show this extreme loss of variation in two populations of Cape mountain zebra relative to free ranging Hartmann's mountain zebra and plains zebra respectively.

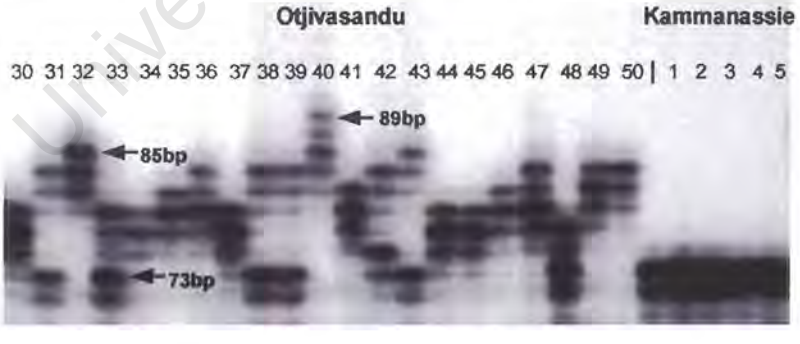

Figure 3.7A. Relative amounts of allelic variation between free ranging Northern Hartmann's mountain zebra (Otjivasandu) and a population of Cape mountain zebra (Kammanassie) at microsatellite locus HTG 5 (Ellegren *et al.* 1992).

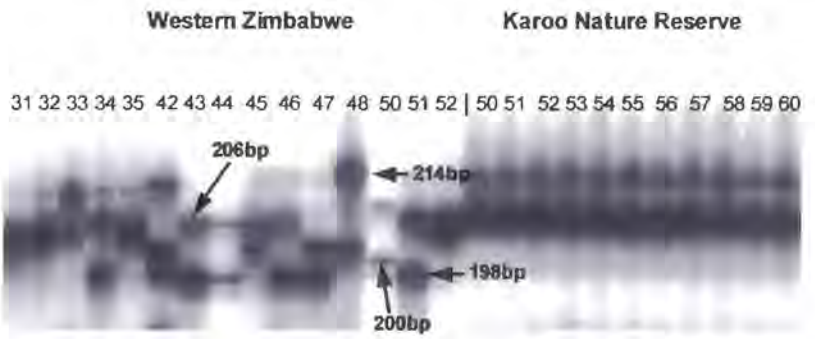

Figure 3.7B. Relative amounts of allelic variation between Zimbabwean plains zebra and a population of Cape mountain zebra (Karoo Nature Reserve) at microsatellite locus LEX 20 (Coogle *et al.* 1996).

3.3.2 Allele frequency histograms

The raw data are best summarised in allele frequency histograms. Included here are four histograms (Fig. 3.8) that illustrate the different population scenarios resulting from random drift in the three small Cape mountain zebra stocks as well as DeHoop, which is a Cape mountain zebra population descended from a mixed origin of equal numbers from Kammanassie and Cradock. Combined data from seven free ranging Hartmann's mountain zebra populations were included for comparison.

It is apparent that the dramatic bottlenecking of Cape mountain zebra populations has promoted a tangible loss in allelic diversity. This loss was further exacerbated by the effect of genetic drift on small populations to the extent that some stocks have become fixed for just a single allele (see also Table 3.2B). Figures 3.8C and 3.8D provide further evidence of continuing drift in the Cradock metapopulation, where one allele is drifting to fixation at the expense of several others that occur at very low frequency. The Hartmann's mountain zebra metapopulation, by contrast, has a greater and generally more evenly spread compliment of alleles. The Cape mountain zebra population at DeHoop Nature Reserve, seeded in the early 1970s, shows a greater allelic diversity than any of the three stock populations at only two of the fifteen studied loci (Fig. 3.8A and 3.8B). At both these two loci, DeHoop possesses alleles that do not occur in other Cape mountain zebra populations, indicating that at the time of seeding, genetic diversity in at least one of its parent populations was higher than it is at present.

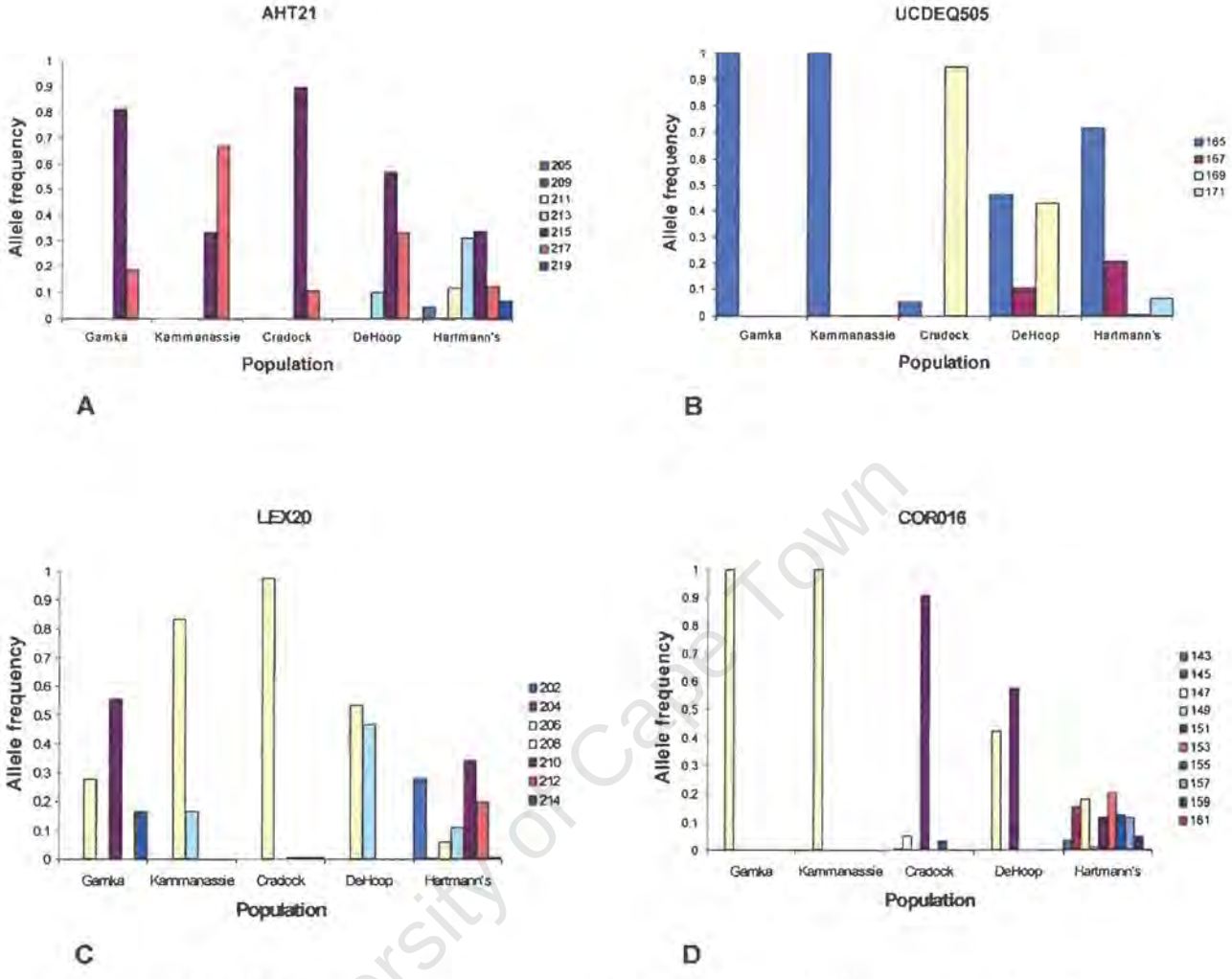

Figure 3.8. Allele frequency histograms for four microsatellite loci showing the distribution of alleles in five groups of mountain zebra; the three relict Cape mountain zebra stocks (Gamka, Kammanassie and Cradock), the mixed-descent Cape mountain zebra population at DeHoop and the entire Hartmann's mountain zebra metapopulation. Fig. A, Locus AHT 21; Fig. B, Locus UCDEQ 505; Fig. C, Locus LEX 20; Fig. D, Locus COR 014.

3.3.3 Inter-locus disequilibrium

The 15 microsatellite loci were tested for linkage disequilibrium using GENEPOP, version 3.1c (Raymond and Rousset, 1995a), using 100 batches of 1000 iterations and 1000 dememorisation steps between each contingency table. The null hypothesis that two loci in a pair-wise comparison are statistically linked was rejected at the 99% level ($p > 0.01$) for all but one pair of loci – HTG 11 and LEX 52 (overall p-value across all populations = 0.008). In tests for linkage disequilibrium, problems may arise when some sample populations are monomorphic at a particular locus. Global pair-wise p-values are calculated from contingency tables according to the χ^2 distribution, with degrees of freedom equal to two times the number of populations where both loci are polymorphic. This situation occurs only four out of fifteen times for the above pair of loci. If the χ^2 value for those four populations happened to be high, the p-value with df = 8 would output a non-significant p-value. This method of summarising the information obtained by Markov chain simulation may sometimes use only a small amount (27% in this case) of the data, thereby producing a misleading result. Furthermore, when either of these loci was omitted from the analyses, general population parameters were not significantly affected. All loci were therefore assumed to be independently assorting.

3.3.4 Nuclear microsatellite genetic variation and population structuring

3.3.4.1 Genetic variation

Genetic variation in Hartmann's mountain zebra was moderate by comparison to values in plains zebra (Section 4.3.3.1) and other microsatellite studies on other large ungulates such as African buffalo (O'Ryan *et al.* 1998), American bison (Wilson and Strobeck, 1999), African wild cat (Wiseman *et al.* 2000) and the northern black rhinoceros (E. Harley pers. comm.). However, these levels were consistent with other taxa such as moose (Broders *et al.* 1999), southern African elephant (Whitehouse and Harley, 2001) and southern black rhinoceros (E. Harley, pers. comm.). As the studied Hartmann's mountain zebra populations are known to be free ranging and outbred, it is proposed that the values in Table 3.2A be considered as baseline values for genetic variation in *E. zebra*.

By contrast, heterozygosity and allelic diversity are significantly lower in all Cape mountain zebra populations (H_O: $p < 0.001$, A: $p < 0.001$; Student's two sample t-test; Table 3.2B), reflecting the trend observed in allele frequency histograms.

Table 3.2. Nuclear genetic variation in mountain zebra (*Equus zebra*) based on 15 microsatellite loci. Two values given for founder number N_F, show that some populations were subjected to two separate bottleneck events. N_S denotes the number of animals sampled in each population; N_A, the number of alleles and N_P is the number of polymorphic loci in each population. The mean number of alleles per locus (A) is corrected for sample size using the jack-knifing procedure. metapop., metapopulation; ?, unknown. H_O, observed proportion of heterozygotes in each population; H_E, unbiased proportion of expected heterozygotes based on allele frequencies calculated according to Nei (1978).

Table 3.2A. Genetic variation in Hartmann's mountain zebra populations

Population	Region	N_F	N_S	N_A	N_P	A	Heterozygosity (H)	
							H_O	H_E
Khomas Hochland	Mid-Central	-	55	69	15	3.44	0.4814	0.5429
Gamsberg	Mid-Central	-	11	55	14	3.36	0.5255	0.5143
Erongo	North Central	-	5	53	14	3.23	0.4333	0.5590
Naukluft	South Central	-	9	49	13	3.16	0.4340	0.4888
Auasberg	Introduced	?	4	36	12	2.40	0.4389	0.4703
Kamanjab	Northern	-	15	58	14	3.39	0.4981	0.5041
Otjivasandu	Northern	-	101	74	15	3.51	0.4716	0.5009
TOTAL			**200**					

Table 3.2B. Genetic variation in Cape mountain zebra populations

Population	Stock	N_F	N_S	N_A	N_P	A	Heterozygosity (H)	
							H_O	H_E
Gamka	Gamka	6	9	31	11	2.07	0.2435	0.2875
Kammanassie	Kammanassie	5	9	28	9	1.78	0.1944	0.2392
Cradock metapop.	Cradock	19	62	47	12	2.03	0.1994	0.2548
Mountain Zebra NP	Cradock	19	12	34	8	2.17	0.1890	0.2760
Karoo NP	Cradock	19 and 76	13	28	9	1.76	0.1470	0.2239
Karoo NR	Cradock	19 and 20	13	29	10	1.85	0.2872	0.2712
Gariep Dam NR	Cradock	19 and 24	10	25	6	1.63	0.1013	0.1590
Bontebok NP	Cradock	19 and 24	14	31	8	1.85	0.2381	0.2326
DeHoop NR	Kammanassie/ Cradock	14	15	33	12	2.11	0.3949	0.3798
TOTAL			**95**					

The level of genetic diversity observed in Auasberg, the only seeded Hartmann's mountain zebra population, is still much higher than in any Cape mountain zebra population. The number of remaining Cape mountain zebras after the recent population bottlenecking of the 1950-1975 period (described in Sections 3.1.6 and 3.1.7) is similar in magnitude to the number of animals relocated to the Auasberg. This comparison suggests that these recent and often cited bottlenecking events were not the only factors leading to the current low levels of genetic variation in extant Cape mountain zebra populations. Instead, results from an Auasberg sample, albeit small, suggest that the genetic diversity was gradually lost from Cape mountain zebra stocks over a longer period of time. This is consistent with the record of hunting and exploitation outlined earlier. It may be that the Auasberg population recovered quickly after its founding, thereby curbing the loss of variation due to drift but a similar demographic recovery took place at MZNP without the restoration of the high levels of genetic diversity observed in Hartmann's mountain zebra. Genetic drift has therefore been affecting gradually decreasing Cape mountain zebra populations for much longer than first imagined.

Allelic diversity (*A*) within Cape mountain zebra populations was severely affected by population bottlenecking with only three populations, the stocks at Gamka and MZNP and the mixed DeHoop, retaining more than a mean of two alleles per locus (Table 3.2B). Observed heterozygosity was also low relative to Hartmann's mountain zebra populations. Gamka was the only aboriginal population to retain a level of H_O greater than 0.2.

The very act of seeding a new population subjects that population to a secondary bottleneck. The relative ease and lesser expense of capturing whole family units as opposed to single individuals from different family groups further exacerbates the effect of the secondary bottleneck, as seeded individuals are already more closely related to each other. The effect of secondary bottlenecking is apparent in levels of allelic diversity but not in observed heterozygosity within Cradock-derived populations. Furthermore, neither allelic diversity nor observed heterozygosity seems to be related to founder number. Both Karoo Nature Reserve (KNR) and BNP, seeded from 20 and 24 individuals respectively, showed higher H_O than their founding stock, with KNR retaining two more polymorphic loci than the MZNP. This may be the result of founding by a more evenly sampled group of individuals or by the common practice of the National Parks Board of intermittently reseeding satellite populations whenever numbers at MZNP or KNP became too high. This served as a boost to the genetic

composition of the seeded population, having the same effect as migration in free ranging populations. The population at KNR resulted from a single translocation event in 1981. By this time 36 animals were moved from MZNP to KNP. This must have been a time when the aboriginal gene pool at MZNP was more diverse than it is at present, leading to the higher H_O and N_P at KNR and a higher N_P at KNP. In these ways, the translocation policy of South African National Parks (SANP), especially the earlier movements, may be justified.

However, the effects of these movements on the aboriginal population seem to be detrimental. Animals were removed from the MZNP as early as the 1960s, but these movements only reached a peak in the 1980s. By the mid-1990s more animals were being translocated from KNP, which had superseded its founding population in size. Both these populations have lost and continue to lose individuals to an increasing number of satellite populations. This loss, coupled with genetic drift owing to decreasing numbers, has resulted in lower estimates of some measures of genetic diversity in MZNP and KNP (the latter seeded with 76 individuals) relative to some satellite populations.

When the entire Cradock metapopulation is considered, H_O and N_P increase whereas allelic diversity decreases relative to the MZNP. This tends to justify further the random movement of animals from MZNP. The Cradock metapopulation contains more genetic diversity than Kammanassie and more polymorphic loci than Gamka although the latter has a higher H_O and a higher allelic diversity. In the future, when new satellite populations are seeded with more than one stock, only those genetically diverse Cradock-derived populations such as MZNP and KNR will contribute individuals. The combined result is therefore not a fair reflection of the potential of the Cradock gene pool as the combined analysis is skewed downward by seeded populations, some with a much lower genetic diversity.

The relatively high level of genetic variation present at Gamka implies that the number of animals present after the last bottleneck event was higher than the estimated three to six. However, if this estimate is correct, the duration of the bottleneck would have to have been short and the population able to re-multiply in a short space of time. There are no demographic data to confirm this. In a 1985 census, the population was stable at 19 individuals and today this number has increased to 28. By contrast, in 1985, the Kammanassie population consisted of 12 individuals and has grown to the present day figure of 38. Although the estimated number of founders of the present population at

Kammanassie was no more than five animals, from the data available, it appears that its bottleneck period was sustained until recently. Genetic drift has thus had a greater effect at Kammanassie, rapidly reducing genetic variation and making it the least diverse of the aboriginal Cape mountain zebra stocks.

DeHoop Nature Reserve, the only Cape mountain zebra population of truly mixed origin, has a higher number of polymorphic loci (N_P) and higher heterozygosity than any other Cape mountain zebra population while allelic diversity (A) is of the same magnitude as that in the most diverse Cape mountain zebra populations (Table 3.2B). This may indicate that Kammanassie and Cradock, the two stocks from which DeHoop was derived, have drifted randomly such that some of their genetic content is different. It is also possible that DeHoop contains unique genetic material as the introduction of animals predates both the Kammanassie bottleneck and the late 1970s when the MZNP population appears to have lost diversity to KNP and KNR. The genetic diversity observed at DeHoop approaches that of the least diverse Hartmann's mountain zebra population, implying that even before recent bottlenecks, historic levels of diversity were not preserved in a combination of only Kammanassie and Cradock genetic elements. Nevertheless, DeHoop is an interesting case study of how best to manage genetic diversity in endangered species as the mixture of two genetically depauperate aboriginal stocks has led to a level of genetic diversity comparable to that of outbred populations. The chance of inbreeding depression via the expression of deleterious recessive mutants is thus much lower.

3.3.4.2 Hardy-Weinberg equilibrium

Of the 225 possible population-locus exact tests for departure from Hardy-Weinberg equilibrium (HWE; Table 3.3), 54 tests could not be computed owing to the presence of monomorphic alleles, 46 of which were associated with Cape mountain zebra populations. Three of the six significant ($p < 0.01$) population-locus exact tests were not significant after sequential Bonferroni correction (Rice, 1989). The null hypothesis of Hardy-Weinberg equilibrium in global tests for each population and each locus was not rejected ($p > 0.05$). We assume therefore that all studied populations are in Hardy-Weinberg equilibrium for all 15 loci.

Table 3.3. Exact tests for departure from Hardy-Weinberg equilibrium (HWE) in mountain zebra populations. P values indicate the probability of rejecting the null hypothesis of random association of alleles, that is HWE. The null hypothesis cannot be rejected at $p > 0.05$ (**) or $p > 0.01$ (*). The null hypothesis of HWE can be rejected for $p < 0.01$. Population/locus combinations still significantly deviating from HWE after sequential Bonferroni correction at $p < 0.01$ were denoted $p < 0.01$B. In some instances no test was possible (-) due to monomorphism across a population. pop, population

Population\ Locus	HTG3	HTG5	HMB 1	HTG7	HTG9	HTG 11	HTG 14	HTG 15	LEX 52	VHL 47	LEX 20	UCDEQ 505	AHT 21	TKY 273	COR 014	Global pop.
Gamka	**	**	*	-	**	**	**	-	**	**	**	-	**	**	-	**
Kammanassie	-	**	**	**	-	*	**	-	-	**	**	-	**	**	-	**
MZNP	**	**	**	**	**	**	**	-	-	-	**	-	-	-	p<0.01	**
Karoo NP	-	**	**	**	**	**	**	-	-	**	-	-	**	**	-	**
Karoo NR.	-	**	**	**	**	**	**	-	-	**	-	**	**	-	**	**
Gariep Dam NR	-	**	-	**	**	**	-	-	-	-	-	-	**	**	-	**
Bontebok NP	-	**	**	**	**	**	**	-	-	**	-	-	-	-	**	**
DeHoop NR	-	**	*	**	**	**	**	-	-	**	*	**	**	**	**	**
Khomas Hochland	*	**	**	**	**	**	**	**	p<0.01	**	p<0.01B	*	**	**	*	**
Gamsberg	**	**	*	**	**	**	**	**	-	**	**	**	**	**	*	**
Auasberg	**	**	**	**	**	**	-	**	-	-	**	**	**	**	**	**
Erongo	**	*	**	**	**	**	**	**	-	**	**	**	**	**	**	**
Naukluft	**	**	**	p<0.01	**	**	**	**	-	-	**	**	**	**	**	**
Kamanjab	**	**	**	**	**	**	-	**	**	**	**	**	**	**	**	**
Otjivasandu	**	**	**	**	**	p<0.01B	**	**	**	**	**	**	**	**	p<0.01B	**
Global locus	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**

3.3.4.3 Analysis of molecular variance (AMOVA)

Several levels of AMOVA were performed to elucidate genetic structure in mountain zebras. The Φ_{ST} index for the various grouping scenarios were tested with 10 000 permutations. The entire mountain zebra metapopulation shows significant genetic structuring (Table 3.4). The two subspecific groups scenario (CMZ vs HMZ) received the greatest support for metapopulation structuring. Genetic structuring was less pronounced, but still significant for the four group scenario when the Cape mountain zebras were divided into two groups (one of Kammanassie, Gamka and DeHoop and the other of MZNP and its derivatives) and when the Hartmann's mountain zebras were divided into the Northern (Kamanjab and Otjivasandu) and Central (Khomas Hochland, Gamsberg, Erongo and Naukluft) regions.

Table 3.4. Analysis of molecular variance in mountain zebra populations. The fixation index Φ_{ST} (equivalent to θ, Weir and Cockerham, 1984) and the percentage of the variation (%Va) that is distributed among groups are given for different grouping scenarios. ***, significant with $p < 0.001$; **, significant with $p < 0.01$; *, significant with $p < 0.05$; **NS**, not significant ($p > 0.05$)

	Combined mountain zebra metapopulation			
	Φ_{ST}		**%Va**	
1 Group	0.2166***		21.66***	
2 Groups	0.2788***		16.74***	
4 Groups	0.2361***		16.55***	
	Cape mountain zebra		**Hartmann's mountain zebra**	
	Φ_{ST}	**%Va**	**Φ_{ST}**	**%Va**
1 Group	0.3368***	33.68***	0.0566***	5.66***
2 Groups	0.4272***	28.43**	0.0611***	1.84^{NS}

When each subspecies was analysed separately using the same groupings, marked genetic structuring was observed among Cape mountain zebra populations and even greater structuring between the two defined Cape mountain zebra groupings, the latter value being almost twice that of its equivalent metapopulation value. Moderate genetic structuring was detected between Hartmann's mountain zebra populations, with marginally greater structuring when the subspecies was split into Northern and Central groups. The high and significant values observed for the combined metapopulation statistics is therefore probably an artifact of the disproportionately large Φ_{ST} values obtained for the Cape mountain zebra metapopulation. This supports the hypothesis that genetic drift and inbreeding have led to the loss of alleles and the genetic

divergence among Cape mountain zebra populations. This is investigated further below.

3.3.4.4 Exact tests of population differentiation

Highly significant allelic heterogeneity was observed in 93 of the 105 pair-wise population comparisons (Table 3.5). Markov chain simulations delivered many χ^2 values of infinity, resulting in p-values of < 0.00001. Less significant values occur between some pair-wise combinations of Cradock-derived Cape mountain zebra populations but even amongst these there were instances of highly significant genetic heterogeneity. Although mostly significant, p-values for exact tests associated with the Auasberg and Erongo Hartmann's mountain zebra populations were larger than for other populations (Table 3.5), with the pair-wise comparisons between the two populations being only one of two non-significant tests; the other non-significant exact test was that of Naukluft (South-Central) vs Kamanjab (Northern). Both results are likely to be caused by the small sample sizes for each of these populations although it is possible that animals from the Erongo Mountains were used to seed the Auasberg population.

Table 3.5. Exact pair-wise tests of differentiation between mountain zebra populations. P values are given for each pair-wise comparison. ***, significant with $p < 0.001$; **, significant with $p < 0.01$; * significant with $p < 0.05$, **NS**, not significant ($p > 0.05$). The box in the top half of the table encloses the MZNP population and its derivatives and the lower right quadrant contains the pair-wise comparisons of Hartmann's mountain zebra populations. Kamma., Kammanassie

	Gamka	Kamma.	MZNP	KNP	KNR	GDNR	BNP	DeHoop	Khomas	Gamsberg	Auasberg	Erongo	Naukluft	Kamanjab
Kammanassie	***													
MZNP	***	***												
Karoo National Park	***	***	***											
Karoo Nature Reserve	***	***	***	*										
Gariep Dam Nature Reserve	***	***	***	**	***									
Bontebok National Park	***	***	*	**	***	***								
DeHoop	***	***	***	***	***	***	***							
Khomas Hochland	***	***	***	***	***	***	***	***						
Gamsberg	***	***	***	***	***	***	***	***	***					
Auasberg	***	***	***	***	***	***	***	***	***	*				
Erongo	***	***	***	***	***	***	***	***	***	**	NS			
Naukluft	***	***	***	***	***	***	***	***	***	***	*	*		
Kamanjab	***	***	***	***	***	***	***	***	***	***	*	**	NS	
Otjivasandu	***	***	***	***	***	***	***	***	***	***	***	***	***	***

3.3.4.5 Pair-wise population structuring

A θ or *Rho* value of 1 indicates complete population subdivision and a value of 0 indicates complete lack of differentiation. Generally, $0.05 < \theta / Rho < 0.1$ implies low to moderate population structuring. $\theta / Rho > 0.1$ suggests moderate to high level population differentiation.

Moderate population genetic structuring is inferred from pair-wise fixation indices in natural populations of Hartmann's mountain zebra ($Rho < 0.0910$, $\theta < 0.0846$) although 7/15 *Rho* values and 3/15 θ values were not significant ($p > 0.05$) after 10 000 permutations (Table 3.6). The highest of these fixation indices are associated with the Erongo population. Given that the Erongo area effectively links the Hartmann's mountain zebra populations of central and northern Namibia, the observed statistics are more likely to be the result of small sample size (n Erongo=5) rather than of genuine population differentiation. The moderate genetic structuring among these natural populations seems indicative of drift owing to isolation by distance or of reduced gene flow due to human encroachment and the erection of fences in the Central region.

The only introduced Hartmann's mountain zebra population, Auasberg, shows higher *Rho* (0.0625 – 0.1796) but moderate θ (0.0472 – 0.0757) relative to the rest of the metapopulation. High values may also be attributed to low sample size (n Auasberg=4) but also to a small number of founders, genetic drift and the inherent high variance associated with R_{ST} and its derivatives.

By comparison, pair-wise fixation indices for Cape mountain zebra (Table 3.6) show grossly high values (*Rho*, 0.3165 – 0.4697; θ, 0.3947 - 0.5289) differentiating the stock populations of Kammanassie, Gamka and the MZNP. Given that the Hartmann's mountain zebra populations reflect the historical degree of population structuring among mountain zebras, it is clear that the marked genetic differentiation among Cape mountain zebra stocks is a result of the severe reduction in the effective population size of each stock. The high differentiation between Kammanassie and the Cradock-derived stocks (*Rho*, 0.4694 – 0.6866; θ, 0.4924 - 0.6514) shows conclusively that although the Kammanassie male "Tom" sired up to ten foals at the MZNP (Novellie, *in litt.*), his contribution to the Cradock gene pool has been negligible. As Tom was one of the first zebras moved from the Kammanassie Mountains in 1970, his genotype over the 15 studied loci may have, by chance, contained more of what are now regarded as Cradock-like alleles.

Pair-wise values between Cradock-derived populations are generally lower (<0.1) and less significant ($p > 0.01$ for 3 out of 12 pair-wise comparisons) for all but those comparisons associated with Gariep Dam Nature Reserve (GDNR). Since each seeded population underwent a secondary bottleneck, these low pair-wise values may appear surprising. However, seeded Cape mountain zebra populations have intermittently received more animals from either the original population at MZNP or from the largest population at KNP. With the exception of KNR (19 years), the time since the last seeding event for any of these populations is no more than four years, much less than one generation. The Cradock-derived populations therefore seem not to have drifted significantly from their aboriginal gene pool or relative to each other. The KNR population has not drifted relative to its progenitor despite being seeded with only 20 individuals and isolated for longer than the other Cradock derivatives. This may be owing to a more random sample from different family groups taken from the MZNP. As these individuals were removed relatively early (1981), it is conceivable that genetic diversity at the MZNP was also greater at this time.

Rho and θ values are high for pair-wise comparisons between GDNR and MZNP and other Cradock-derived populations. Genetic diversity at GDNR is the lowest observed for any mountain zebra population, including BNP and KNR which were seeded with the same number or fewer founders, respectively. This highlights the random nature of genetic drift.

Being mixed at a ratio of 50:50, DeHoop was expected to show intermediate levels of genetic differentiation relative to its founding stocks. In contrast to this, DeHoop is highly differentiated from Kammanassie and Gamka but does show intermediate structuring with MZNP. The Kammanassie population undoubtedly contributed to the observed high genetic diversity of DeHoop. However, the presence of rare alleles at DeHoop (see Fig. 3.8) that have subsequently been lost in Kammanassie owing to drift, significantly increases the amount of pair-wise differentiation between the two populations.

Table 3.6. Pair-wise fixation indices between mountain zebra populations. Values for the F_{ST} estimator θ are given above the diagonal and values for the R_{ST} estimator *Rho* are given below the diagonal. **NS**, not significant ($p > 0.05$); *, significant at $p < 0.05$; **, significant at $p < 0.01$; after 10 000 permutations. Upper left quadrant shows pair-wise comparisons of Cape mountain zebra populations. The box inside the upper left quadrant encloses the MZNP population and its derivatives and the lower right quadrant contains the pair-wise comparisons of Hartmann's mountain zebra populations. Kamma., Kammanassie

	Gamka	Kamma	MZNP	KNP	KNR	GDNR	BNP	DeHoop	Khomas	Gamsberg	Auasberg	Erongo	Naukluft	Kamanjab	Otjivasandu
Gamka		0.3947 **	0.4932 **	0.5643 **	0.5108 **	0.6285 **	0.5477 **	0.3506 **	0.2595 **	0.2960 **	0.3795 **	0.2908 **	0.2853 **	0.2476 **	0.2878 **
Kammanassie	0.4480 **		0.5289 **	0.5934 **	0.4924 **	0.6514 **	0.5555 **	0.3021 **	0.2917 **	0.3631 **	0.4326 **	0.3557 **	0.3438 **	0.3562 **	0.3210 **
MZNP	0.3165 **	0.4697 **		0.0641 *	0.0808 **	0.1797 **	0.0071 NS	0.2222 **	0.2577 **	0.2736 **	0.4111 **	0.3366 **	0.3395 **	0.3346 **	0.2847 **
Karoo National Park	0.4125 **	0.6054 **	0.0486 NS		0.0519 *	0.0935 *	0.0504 NS	0.2758 **	0.2948 **	0.3330 **	0.4690 **	0.4248 **	0.4203 **	0.3902 **	0.3142 **
Karoo Nature Reserve	0.3540 **	0.5255 **	0.0269 NS	0.0232 NS		0.1722 **	0.0676 **	0.1944 **	0.2592 **	0.3069 **	0.4258 **	0.3614 **	0.3479 **	0.3430 **	0.2878 **
Gariep Dam Nature Reserve	0.4995 **	0.6866 **	0.1527 *	0.1045 **	0.1378 **		0.1927 **	0.2720 **	0.3071 **	0.3531 **	0.5212 **	0.4517 **	0.4500 **	0.4157 **	0.3314 **
Bontebok National Park	0.4332 **	0.5899 **	0.0799 NS	0.0626 **	0.0297 NS	0.1954 **		0.2424 **	0.2751 **	0.3202 **	0.4656 **	0.3996 **	0.3932 **	0.3776 **	0.2995 **
DeHoop	0.3849 **	0.4890 **	0.1208 **	0.2122 **	0.1305 **	0.2074 **	0.2278 **		0.1697 **	0.1815 **	0.2631 **	0.1961 **	0.2011 **	0.2304 **	0.2038 **
Khomas Hochland	0.3833 **	0.5016 **	0.2716 **	0.2975 **	0.2745 **	0.3426 **	0.3051 **	0.2192 **		0.0329 **	0.0726 **	0.0846 **	0.0435 **	0.0677 **	0.0723 **
Gamsberg	0.3799 **	0.4785 **	0.2305 **	0.2391 **	0.2075 **	0.2987 **	0.2439 **	0.1538 **	0.0284 NS		0.0628 **	0.0465 *	0.0593 **	0.0527 **	0.0460 **
Auasberg	0.4647 **	0.5732 **	0.3245 *	0.3111 **	0.2946 **	0.3967 **	0.3322 **	0.2923 **	0.1511 **	0.1056 *		0.0525 NS	0.0472 NS	0.0757 **	0.0714 **
Erongo	0.3514 **	0.3759 **	0.1279 NS	0.1147 NS	0.1232 *	0.1246 NS	0.1513 *	0.0317 NS	0.0848 *	0.0382 NS	0.0625 NS		0.0056 NS	0.0258 NS	0.0347 *
Naukluft	0.3101 **	0.5071 **	0.2479 **	0.2535 **	0.2236 **	0.3137 **	0.2829 **	0.2039 **	0.0526 *	0.0331 NS	0.1337 NS	0.0795 NS		0.0157 NS	0.0449 **
Kamanjab	0.2833 **	0.4699 **	0.2351 **	0.2159 **	0.2199 **	0.3009 **	0.2584 **	0.2204 **	0.0420 **	0.0414 *	0.1479 **	0.0910 NS	0.0271 NS		0.0224 **
Otjivasandu	0.4029 **	0.5406 **	0.2692 **	0.2675 **	0.2779 **	0.3474 **	0.3048 **	0.2438 **	0.0497 **	0.0210 NS	0.1796 **	0.0612 *	0.0586 **	0.0209 *	

When each Cape mountain zebra population was compared to the entire Hartmann's mountain zebra metapopulation (Table 3.7), intermediate values usually resulted. This may seem counter-intuitive, as it appears that each Cape mountain zebra stock is more closely related to Hartmann's mountain zebras than to each other. This is precisely the case, but this marked observed differentiation among Cape mountain zebra stocks and between these and the Hartmann's mountain zebra metapopulation was not the result of a natural process but of a prolonged history of human-mediated population decimation. Individual populations within the Cape mountain zebra metapopulation may contain a high number of private alleles relative to each other but contain very few relative to the Hartmann's mountain zebra metapopulation. They therefore appear more closely related to Hartmann's mountain zebra than to each other.

Table 3.7. The F_{ST} estimator θ and the R_{ST} estimator *Rho* between Cape mountain zebra populations and the Hartmann's mountain zebra metapopulation. NS, not significant ($p > 0.05$); *, significant at $p < 0.05$; **, significant at $p < 0.01$; after 10 000 permutations.

	Hartmann's mountain zebra metapopulation	
	Rho	θ
Gamka	0.36689 **	0.19306 **
Kammanassie	0.50338 **	0.28509 **
MZNP	0.24819 **	0.18254 **
Karoo National Park	0.24910 **	0.28761 **
Karoo Nature Reserve	0.23394 **	0.26012 **
Gariep Dam Nature Reserve	0.31493 **	0.29234 **
Bontebok National Park	0.27649 **	0.27127 **
DeHoop	0.20651 **	0.15864 **

When the Cape and Hartmann's mountain zebra metapopulations are analyzed in their entirety, the resultant pair-wise values (*Rho* = 0.1965, θ = 0.1948) are much lower than those between each Cape mountain zebra stock (Table 3.6) and between these and all Hartmann's mountain zebras (Table 3.7). This lower value indicates that Cape mountain zebras have retained much of their historic genetic variation, albeit in three separate, relict and highly inbred stocks.

3.3.4.6 Genetic distance

Genetic distances among Cape mountain zebra populations, like fixation indices, suggest population bottlenecking, followed by drift. All three Cape mountain zebra stock populations are very divergent from each other (Figs 3.9A, B and C). MZNP and the populations seeded from it form a nested group with DeHoop linked distally. The branch lengths for the Cape mountain zebra stocks show that each is more genetically different to the other than either is to any Hartmann's mountain zebra population.

D_S and D_C (Figs 3.9A and 3.9B) resolve a degree of phylogeographic structuring between mid-Central and northern Hartmann's mountain zebra but this only receives bootstrap support for D_C. The $(\delta\mu)^2$ distance statistic succeeds in separating Cape mountain zebras into their respective stocks and from Hartmann's mountain zebra but provides bootstrap support for only the node separating Kammanassie from Gamka. The position of the North-Central Erongo population varies with distance statistic. Only D_S places it correctly between the northern and mid-Central Hartmann's mountain zebras but this too is not significant (Fig. 3.9A). This variation is probably owing to small sample size. That sample size affects genetic distance is apparent in all three neighbour-joining trees. Those Hartmann's mountain zebra populations with the smallest sample size (Auasberg and Erongo) show the deepest branch lengths. This phenomenon was reported for $(\delta\mu)^2$ in a study of humpback whale populations (Valsecchi *et al.* 1997). Branch lengths are generally longer for D_C suggesting that this statistic considers more within-population variation. The South-Central Naukluft population groups more closely with the mid-central group for D_C and $(\delta\mu)^2$ but neither association is significant. The close association between Naukluft and Auasberg for all distance statistics may suggest the origin of Auasberg founders. From these data we may conclude that the majority of the structure observed among Hartmann's mountain zebra populations is generated between the strong affiliations of the northern Otjivasandu and Kamanjab and the mid-central Khomas Hochland and Gamsberg. The South-Central and North-Central populations show no significant genetic partitioning.

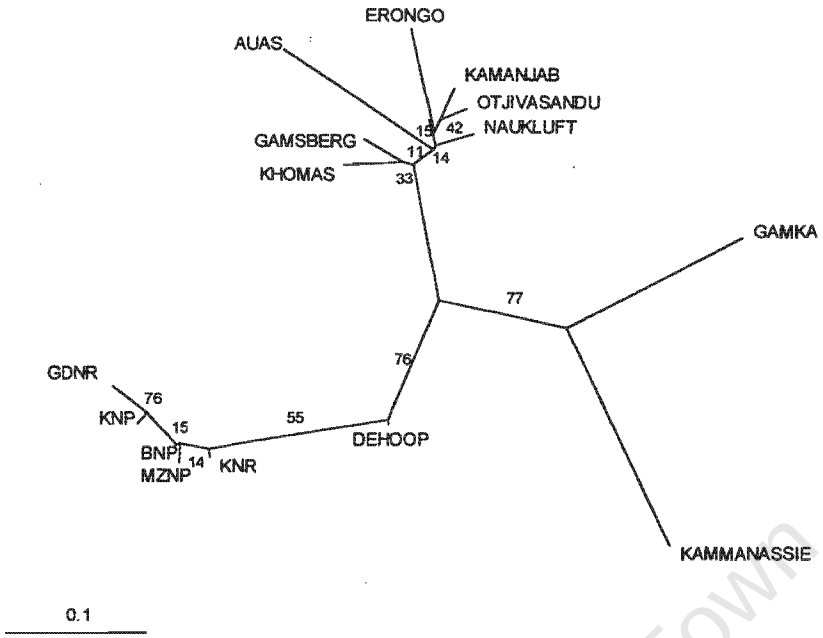

Figure 3.9A. D_S, standard genetic distance, Nei (1972)

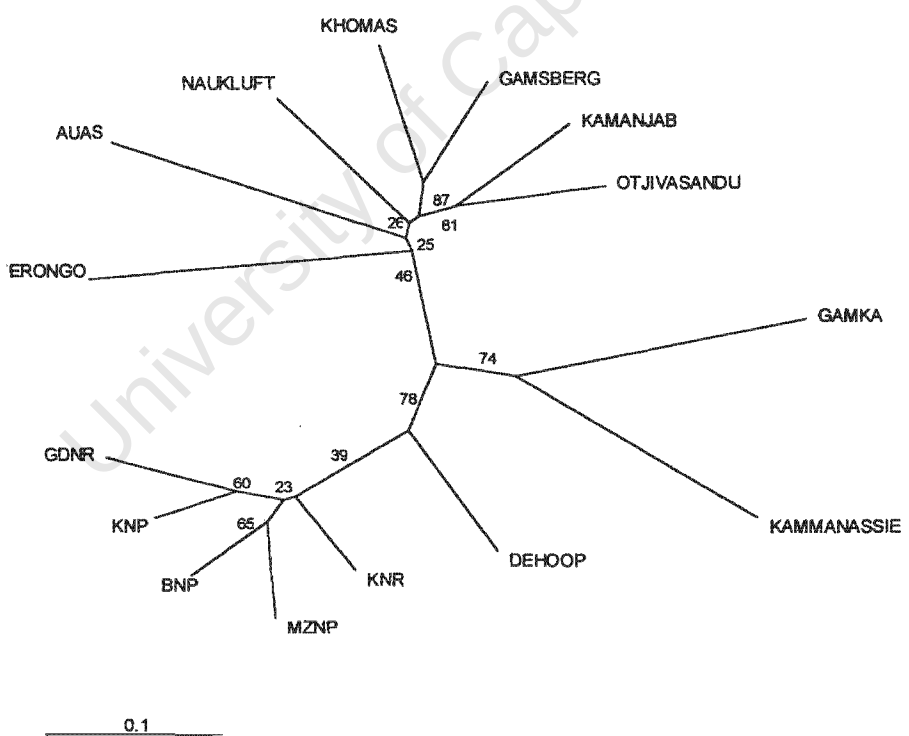

Figure 3.9B. D_C, Cavalli-Sforza and Edwards (1967)

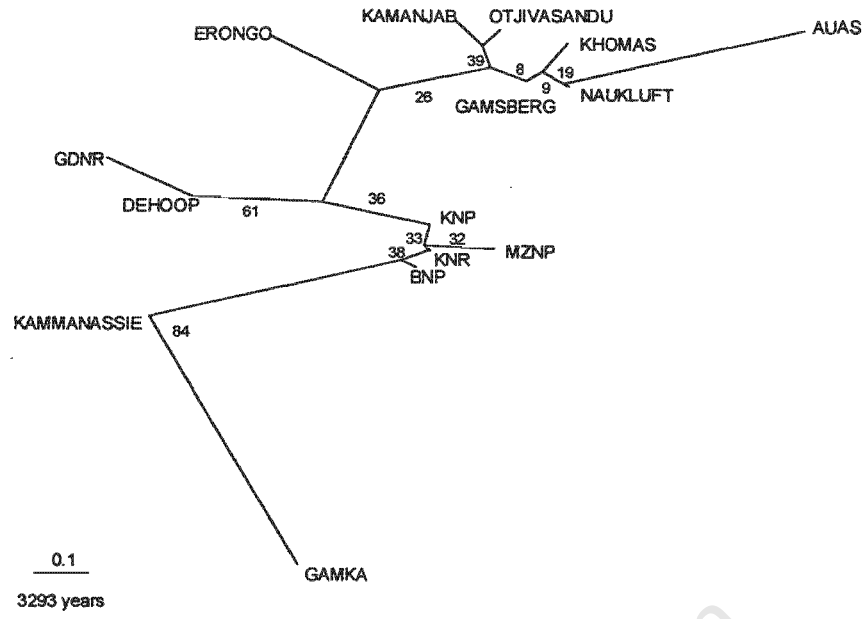

Figure 3.9C. $(\delta\mu)^2$, Goldstein *et al.* (1995)

Figure 3.9. Unrooted neighbour joining trees of three genetic distance measures between populations of mountain zebras. Nodal significance was determined by 1000 bootstrap replicates.

3.3.4.7 Isolation by distance

Genetic distances between free ranging Namibian mountain zebras are likely to result from gene flow as most population numbers are large enough to assume mutation-drift equilibrium. To avoid confounding errors, all Cradock-derived populations were grouped together and the introduced Auasberg population was excluded from the analysis of isolation by distance. D_C performed best with the highest coefficient of regression, followed by D_S and $(\delta\mu)^2$ (Fig. 3.10, Table 3.8). The variance associated with $(\delta\mu)^2$ is apparent in the scatter of data points of Fig. 3.10C, despite the relatively high number of loci used in this study. All three distance measures show significant isolation by distance ($p < 0.01$, Table 3.8) as may be expected from a species distributed one-dimensionally along linear mountain ranges. No two populations were separated by a geographic distance in the range of 700 – 1200 km. To test further for isolation by distance at this level, populations in the Huns Mountains and the Fish River Canyon need to be sampled. Although considered to be Hartmann's mountain zebras, these populations occur almost midway between the sampled populations in this study. The population density of mountain zebras in this region is low, owing to uncontrolled hunting and lack of suitable watering points. Sampling mountain zebras from this area is likely to be very difficult.

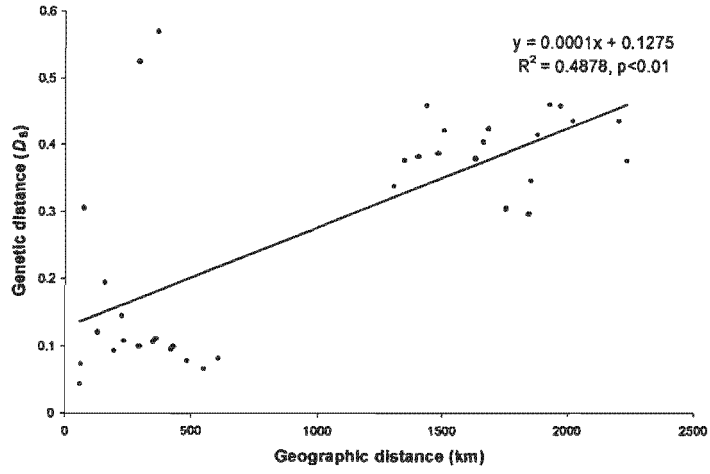

Figure 3.10A. D_S, standard genetic distance, Nei (1972)

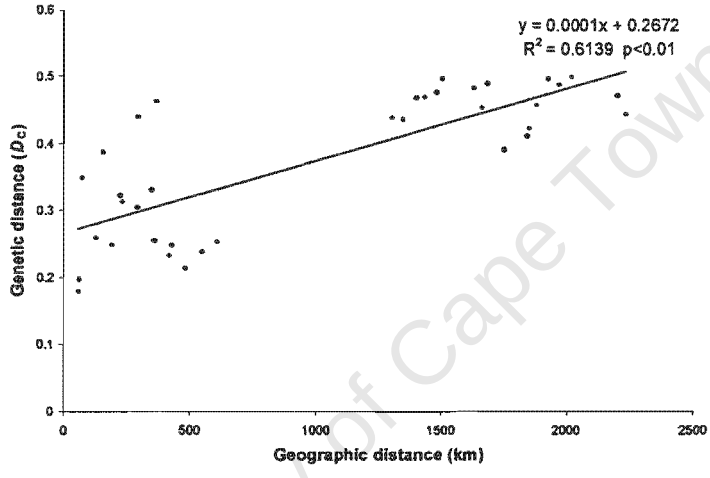

Figure 3.10B. D_C, Cavalli-Sforza and Edwards (1967)

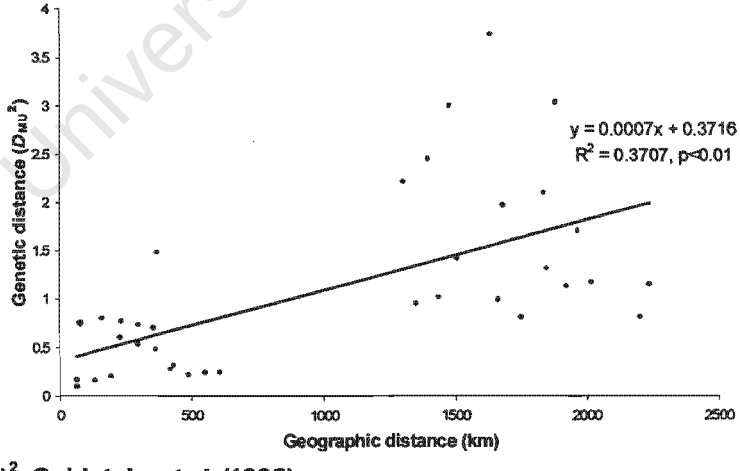

Figure 3.10C. $(\delta\mu)^2$, Goldstein *et al.* (1995)

Figure 3.10. Correlation of geographic distance with three measures of genetic distance in mountain zebra populations.

Table 3.8. Test for isolation by distance in mountain zebras using three distance statistics (D_S, D_C and $(\delta\mu)^2$).

	Coefficient of regression (R^2)	P (any correlation > observed correlation)	Slope of regression line	Y-intercept
D_S	0.4878	0.00200	0.00014858	0.1274873
D_C	0.6139	0.00270	0.00010210	0.2766853
$(\delta\mu)^2$	0.3707	0.00560	0.00072456	0.3716013

Within-population genetic diversity (H_E) was negatively correlated with all measures of genetic distance (Table 3.9). The coefficient of regression (R^2), measuring the strength of the correlation, was highest and most significant for D_S. This implies that a high percentage of the variation in D_S is generated by the biasing effect of genetic diversity. D_C is less affected and the variation in $(\delta\mu)^2$ certainly contributes to a low observed R^2. As all three measures of genetic distance are thought to still be linear at this level, the correlation may be confounded by biologically-driven factors such as drift which will decrease genetic diversity and drive differentiation between small populations. The comparative influence of the various distances on H_E is, however, a qualitative reflection of a real trend.

Table 3.9. Correlation of within-population genetic diversity with three genetic distance statistics (D_S, D_C and $(\delta\mu)^2$).

	Coefficient of regression (R^2)	Significance	Slope of regression line	Y-intercept	X-intercept H_E at distance=0
D_S	0.7695	0.00710	-1.71364228	1.0188679	0.5945628
D_C	0.5150	0.01400	-0.90844081	0.7693498	0.8468904
$(\delta\mu)^2$	0.1799	0.03840	-4.63430508	3.1089467	0.6708550

3.3.4.8 Divergence times

The gross estimates of population subdivision and genetic distance in Cape mountain zebra can be attributed to relatively recent anthropogenic factors. The calculation of divergence times between Cape mountain zebra stocks and Hartmann's mountain zebra will not correspond to the known demographic history of mountain zebras since accelerated drift will upwardly bias coalescence estimates. The relatively low fixation index comparing the two metapopulations (Section 3.3.4.5) suggests that the Cape mountain zebra metapopulation still retains significant historic heterogeneity. If it is assumed that the historic genetic compliment of Cape mountain zebra populations is roughly approximated by the current genetic compliment of the Cape mountain zebra metapopulation, then the divergence between one population containing all Cape

mountain zebras and each naturally occurring Hartmann's mountain zebra population may be calculated. This assumption may be biased by the fact that all Cape mountain zebra populations have been reduced to very small numbers and subjected to genetic drift such that some alleles may have been collectively lost to the metapopulation. The calculated divergence times may therefore overestimate the actual time when two populations split.

Table 3.10. Divergence times in years with 95% confidence limits, assuming a generation time of 16.5 years, dating the time of split between mountain zebra populations, based on the $(\delta\mu)^2$ distance statistic. The upper left box displays divergence times between Cape mountain zebra stocks and the middle box shows times between the Cape mountain zebra metapopulation and each free-ranging Hartmann's mountain zebra population. Kamma., Kammanassie; CMZ, Cape mountain zebra; Gams., Gamsberg

	Gamka	Kamma.	Cradock	All CMZ	Khomas	Gams.	Erongo	Naukluft	Kamanjab
Kammanassie	30706 ±10906								
Cradock	60205 ±15534	29740 ±12757							
Khomas	122060 ±70749	58152 ±28412	54450 ±38232	49299 ±37668					
Gamsberg	99241 ±47568	40848 ±15655	33443 ±21732	29217 ±21088	4145 ±2455				
Erongo	168944 ±73848	92641 ±49500	55738 ±27406	54893 ±29378	41049 ±24509	32517 ±23663			
Naukluft	90629 ±54732	38513 ±18995	40928 ±26440	34851 ±25676	7767 ±2495	6318 ±2294	37910 ±19196		
Kamanjab	85237 ±39037	45637 ±17023	33121 ±13844	30384 ±12395	18834 ±9739	12878 ±6520	40445 ±21007	9578 ±3059	
Otjivasandu	123830 ±59762	68737 ±27124	47045 ±26360	45596 ±25957	10383 ±5151	8813 ±4427	29056 ±12033	9860 ±4829	6801 ±2817

The calculated divergence times provide an approximation of how long each of these populations has been separated. All estimates are within the range where $(\delta\mu)^2$ is assumed to be linear (Paetkau *et al.* 1997) although the large variance of this statistic (see Fig. 3.10C) further stresses that these values should be considered as rough estimates only.

The gross population differentiation already observed (Table 3.6) contributes to unrealistically high divergence times between Cape mountain zebra stocks and between these and each Hartmann's mountain zebra population. This was circumvented to some extent by combining all the Cape mountain zebra data. The values are then consistent with what was expected for divergence between populations in the north and south of the historic distribution. In evolutionary terms this time span is

too short for processes such as speciation to occur, especially given that the range of mountain zebras was continuous until human intervention, no more than 350 years ago. The divergence observed must therefore be due to isolation by distance.

3.3.4.9 Gene flow

Stock Cape mountain zebra populations have been isolated from each other for more than 100 years. During this time, they rapidly lost genetic diversity owing to a reduction in numbers and by a subsequent acceleration in genetic drift. The results above show that the consequence of this accelerated drift is marked pair-wise genetic differentiation. Genetic estimates of the effective number of migrants between Cape mountain zebra populations is unlikely to be of any practical significance as populations have not exchanged genetic material and drift alone would be the cause of population differentiation. To avoid spurious results such as those observed for each Cape mountain zebra stock in Table 3.10, the results reported here are the calculated N_em values between the Cape mountain zebra metapopulation (pooled stocks) and each free-ranging Hartmann's mountain zebra population.

Table 3.11. Effective number of migrants (N_em) between the Cape mountain zebra metapopulation and naturally occurring Hartmann's mountain zebra populations. Values for the N_em estimate obtained from *Rho* (N_em_{RST}) according to Ciofi and Bruford (1999) are given below the diagonal and values obtained by the private alleles method (N_em_{PA}) (Slatkin, 1985) are given above the diagonal. metapop., metapopulation

	CMZ metapop.	Khomas Hochland	Gamsberg	Erongo	Naukluft	Kamanjab	Otjivasandu
CMZ metapop.		0.4086	0.3488	0.2696	0.2679	0.2457	0.2542
Khomas Hochland	0.4435		1.0459	0.3462	1.1523	0.7744	1.0722
Gamsberg	0.6524	4.2730		0.7185	1.1022	1.6391	1.4614
Erongo	1.1983	1.3491	3.1444		0.7174	0.7783	0.6137
Naukluft	0.6081	2.2502	3.6446	1.4467		1.7638	0.8550
Kamanjab	0.5843	2.8528	2.8948	1.2493	4.4903		1.9233
Otjivasandu	0.4270	2.3898	5.8334	1.9180	2.0063	5.8695	

Both methods of calculating gene flow returned very similar values for the number of migrants among all populations per generation (N_em_{PA} = 0.9391, N_em_{RST} = 0.9403). Assuming a generation time of 16.5 years, one migrant would have been exchanged among mountain zebra populations every 17.5 to 17.6 years. Excluding the South African metapopulation, very different estimates of gene flow were estimated among free ranging Hartmann's mountain zebra populations in Northern and Central Namibia (N_em_{PA} = 0.7440, N_em_{RST} = 12.0322). Discrepancies between the two estimators of gene flow were also observed at the between-population level (Table 3.11). Values obtained for N_em_{RST} were in the range of 1.09 – 4.44 times higher than the private allele estimates, with those pairs of population with the largest sample size difference producing the particularly discrepant estimates. In populations with sample size inequality, the size of a fixed allele at a locus in the small population may be skewed relative to the mean allele size at the same locus in the larger population. Depending on the amount of skew, R_{ST} and thus N_em_{RST} may vary significantly. Both estimates resulted in gene flow values that covaried significantly with each other (Mantel test R^2 = 0.872, $p < 0.01$), indicating that the relative amounts of between-population gene flow detected by both methods were similar.

3.3.4.10 Population assignment tests

Assignment tests provided support for structure in Hartmann's mountain zebra populations. Locally, 94% of individuals are assigned correctly to the Central and 94% to the Northern region. Within the Central region, there is good support (94%) for the mid-Central grouping but moderate support for South-Central and North-Central sub-regions. This moderate support, as for fixation indices and N_em, is probably owing to high gene flow (Table 3.11) between these essentially contiguous distribution ranges.

Cape mountain zebra populations show a higher proportion of correctly assigned individuals owing to the fact that genetic drift has markedly differentiated allele frequencies amongst the three stocks. Within the Cradock metapopulation, values are lower since these populations are all derived from the same stock.

Table 3.12. Percentage of multilocus mountain zebra genotypes correctly assigned to their populations or stocks/regions of origin. Cape mountain zebra results are tabulated on the left and Hartmann's mountain zebra on the right. Stocks/Regions and corresponding values are written in bold.

Population/Stock	% of individuals correctly assigned		Population/Region	% of individuals correctly assigned	
	Population	**Stock**		Population	**Region**
Gamka		**89**	**Central**		**96**
Kammanassie		**100**	**Mid-Central**		**94**
Cradock		**98**	Khomas Hochland	89	
MZNP	67		Gamsberg	91	
Karoo NP	85		**South-Central**		**89**
Karoo NR	92		Naukluft	89	
Gariep Dam NR	100		**North-Central**		**80**
Bontebok NP	93		Erongo	80	
Kammanassie/ Cradock			**Northern**		**94**
DeHoop	93		Kamanjab	93	
			Otjivasandu	78	
			Unknown		
			Auas	100	

3.3.4.11 Principal component (PC) analysis

Significance tests using 1000 permutations showed that PC one was significant (55.9% inertia $p < 0.01$, abscissa axis, Fig. 3.11) whereas PC two was not (16.3% inertia $p > 0.1$, ordinate axis, Fig. 3.11). For ease of interpretation, the data were plotted against these first two principal components (Fig. 3.11) which account for 72.1% of the inertia in the mountain zebra data set. From this, moderate regional structuring between Northern and Central Hartmann's mountain zebra populations is apparent but intra-regional structuring within the Central region is not. The marked differentiation that has thus far characterised the three Cape mountain zebra stocks is evident and, unlike what was found in previous analyses, DeHoop appears to lie intermediate between its founding populations.

Figure 3.11. Mountain zebra populations subjected to principal component analysis. Populations are plotted against the first two principal components which account for 72.1% of the variance in the data. Namibian populations from the Northern region are colour-coded in pink, the Central region in red and Cape mountain zebra populations in blue.

3.3.4.11 Effective population size

N_e estimates from Cape mountain zebra genetic data are much higher than direct population estimates, representing a clear shift from mutation-drift equilibrium in these populations (Table 3.13). This can be explained by the fact that extant Cape mountain zebra populations were reduced to very small numbers. Although some loci are already monomorphic in some Cape mountain zebra populations, indicating prolonged drift, those loci which are polymorphic still contain enough remnant variation to maintain a significantly high N_e. Thus, Cape mountain zebra stocks retain more gene diversity than their small sizes imply. The N_e values presented in Table 3.13 therefore give an indication of the population size at which the current levels of within-population genetic diversity can be maintained. It is also imperative that these numbers be reached as soon as possible since as long as the actual population size remains lower than N_e, alleles will continue to be lost. When dealing with endangered taxa that have experienced similar population bottlenecks, it is proposed that, instead of drawing erroneous conclusions from direct estimates of N_e, conservation authorities regard genetic estimates as the "minimum effective population size" required to prevent further loss of genetic diversity. In the case of the two critically small Cape mountain zebra stocks located at Gamka and Kammanassie, both require an N_e of almost an order of

magnitude higher than their current effective population size to maintain current levels of genetic diversity. This is clearly a grave problem for Western Cape Nature Conservation, the body that manages these populations. The increased genetic diversity present in the DeHoop population translates to an effective population size much higher than that of other Cape mountain zebra stocks, emphasizing that genetic diversity can only be maintained in herds of mixed origin and adequate size.

As all Hartmann's mountain zebra populations show relatively high levels of expected heterozygosity (H_E), values for N_e are accordingly high. Genetic estimates that assume the SMM consistently outperform those that assume the IAM in approximating direct estimates of N_e in Hartmann' s mountain zebra. Only in the Erongo population are both genetic estimates of N_e for a Hartmann's mountain zebra population higher than the direct estimate. This again may be owing to small sample size as H_E is highest in this population. If small sample size is corrected for by using the lowest H_E value in free ranging Hartmann's mountain zebra populations, genetic estimates of N_e would still be higher than the direct estimate. This implies that either the census data reported by Novellie *et al.* (2002) are flawed or that the Erongo population has experienced a reduction in numbers recently enough to maintain a heterozygote excess.

Table 3.13. Estimates of effective population size N_e in mountain zebras. a, T. Barry (pers. comm.); b, G. Cleaver (pers. comm.); c, W. Dorgeloh (pers. comm.); d, Novellie *et al.* (2002); e, W. Dorgeloh (pers. comm.); f, P. Lloyd (pers. comm.); g, assuming 78% of populations consist of breeding adults, a sex ratio of 1:1.7 (Dorgeloh pers. comm.) and $1/N_e=1/4$ (1/Br M +1/Br F). SMM, step-wise mutation model; IAM, infinite alleles model; no est., no estimate obtained; Nth, North; Sth, South; In parentheses, summed genetic estimates of N_e of a particular locality for direct comparison with census and direct estimates.

Population		Extant Population Size	Effective Population size (N_e) Direct Estimate	Assuming SMM	Assuming IAM
Gamka Mountain Nature Reserve		28 [a]	13 [a]	534	450
Kammanassie Nature Reserve		38 [b]	29 [b]	406	355
Mountain Zebra National Park		218 [c]	159 [e]	511	433
DeHoop Nature Reserve		54 [d]	24 [f]	910	706
Karoo National Park		250 [d]	182 [g]	380	334
Karoo Nature Reserve		120 [d]	87 [g]	506	430
Gariep Dam Nature Reserve		58 [d]	42 [g]	234	215
Bontebok National Park		17 [d]	12 [g]	404	353
Khomas Hochland	Mid-Central	7868 [d]	5722 [g]	2245	1419
Gamsberg				1739 (5123)	1174 (3439)
Auasberg		no est.	no est.	1139	846
Erongo	Nth-Central	1135 [d]	826 [g]	1850	1230
Naukluft	Sth-Central	2338 [d]	1700 [g]	1470	1033
Kamanjab				1708	1158
Otjivasandu	Northern	6730 [d]	4895 [g]	1814 (3522)	1212 (2370)

3.3.4.12 Population bottlenecking

Heeding the 0.81 cut-off value, Garza and Williamson's *M* statistic (2001) approximated the demographic history of most Cape mountain zebra populations (Table 3.14). The highest value corresponded to DeHoop. Given results from previous analyses, Kammanassie and GDNR showed unusually high *M*, both values above 0.81. However, if minimum *M* is calculated by subtracting the variance in *M* from the given average value, then only DeHoop, the Cape mountain zebra population likely to consist of a wide range of allele sizes, makes the cut-off, with an *M* of exactly 0.81.

In Hartmann's mountain zebra, minimum *M* was below 0.81 in the Erongo, Naukluft and Auasberg populations. Censuses of the Erongo and Naukluft Mountains (Novellie *et al.* 2002) enumerated these populations at over 1135 and 2338 respectively (Table 3.13). The Naukluft area is remote and sparsely populated and it is more likely that the low *M* is due to insufficient sample size than to recent population bottlenecking. The same may be concluded for Auasberg although this population does have a recent history of population bottlenecking. An increase in game farming in the Karibib-Omaruru districts may be applying greater pressure to the zebras of the Erongo Mountains as two-metre high game fences are less breakable than traditional cattle fences. Furthermore, despite permits being required to hunt mountain zebras, farmers in these remote areas are able to justify the illegal shooting of zebras in the dryer months, when cattle and sheep are often forced away from critical watering points. From the genetic data, it is apparent that mountain zebras are slowly disappearing from the Erongo Mountains.

Minimum values also suggest that population bottlenecking has affected the allele size distribution of the entire Cape mountain zebra metapopulation but not that of the Hartmann's mountain zebra metapopulation (Table 3.14).

If heterozygote excess were to be manifested in any genetically bottlenecked organism, the demographic histories of Cape mountain zebra populations suggest that the phenomenon is most likely to be found in this subspecies. Yet, the method of Luikart *et al.* (1998) detected significant ($p < 0.05$) heterozygote excess only in the DeHoop population, where due to their recent population history, most individuals are heterozygous. The Khomas Hochland ($p < 0.01$), Otjivasandu ($p < 0.05$) and the entire Hartmann's mountain zebra metapopulation ($p < 0.05$) showed significant heterozygote excess under the assumption of IAM, none of which is consistent with demographic records.

Table 3.14. Comparison of two tests for population bottlenecking in all mountain zebra populations and both subspecific metapopulations. $M < 0.81$ (Garza and Williamson, 2001) or p-values for heterozygote excess < 0.05 (Luikart *et al.* 1998) indicates population bottlenecking. Minimum *M* calculated as *M* minus variance of *M*.

Population	Garza & Williamson's *M* (minimum *M*)	Variance of *M*	Number of non-monomorphic loci N_P	P-value: Heterozygote excess SMM	IAM
Gamka	0.795 (0.733)	0.062	11	0.6499	0.2324
Kammanassie	0.839 (0.770)	0.069	9	0.5000	0.0645
Mountain Zebra NP	0.748 (0.693)	0.055	8	0.6797	0.3711
DeHoop Nature Reserve	0.860 (0.810)	0.050	12	0.0386	0.0053
Karoo National Park	0.783 (0.708)	0.075	9	0.5450	0.1797
Karoo Nature Reserve	0.786 (0.722)	0.064	10	0.2158	0.0654
Gariep Dam NR	0.833 (0.755)	0.078	6	0.9453	0.5781
Bontebok National Park	0.798 (0.738)	0.060	8	0.7266	0.1250
Khomas Hochland	0.919 (0.891)	0.028	15	0.7193	0.0011
Gamsberg	0.911 (0.882)	0.029	14	0.5242	0.1206
Auasberg	0.811 (0.760)	0.051	12	0.3667	0.2349
Erongo	0.747 (0.692)	0.055	14	0.8917	0.5242
Naukluft	0.843 (0.804)	0.039	13	0.6066	0.1527
Kamanjab	0.858 (0.815)	0.043	14	0.7492	0.0969
Otjivasandu	0.860 (0.824)	0.036	15	0.7003	0.0128
All Cape	0.832 (0.781)	0.051	14	0.9973	0.3129
All Hartmann's	0.911 (0.888)	0.023	15	0.9872	0.0151

3.3.5 Mitochondrial haplotype structuring

3.3.5.1 Sequence variation

The 23 mountain zebra control region lineages consisted of 14 haplotypes. Of the 781 bp sequenced, 12.4% of the sites (97 sites) were variable when the *E. quagga* outgroup was included in the data set and 7.4% (58 sites) were variable when the outgroup was excluded. Exclusion of sites with alignment gaps or missing data revealed 52 site patterns and 43 polymorphic sites within the 14 *E. zebra* haplotypes, with a transition/transversion ratio of 11.96 (Table 3.15). Four haplotypes were shared among 13 individuals from seven populations, with four individuals (representing three populations) sharing the most common haplotype. Each Cape mountain zebra stock consisted of a unique haplotype and both Kammanassie and Cradock haplotypes are present in the DeHoop population, reflecting its mixed origin. The *E. zebra* control region was characterised by high (0.921 ± 0.033) haplotype diversity. Substitution rate heterogeneity was also high (α = 0.11) but comparable to that obtained by Oakenfull *et al.* (2000) for the genus. Gamma corrected nucleotide diversity as calculated by Kimura's (1980) 2-parameter method π^1 (0.01630 ± 0.00649) and by maximum likelihood π^2 (0.01942 ± 0.00776) was low, inferring a recent coalescence of *E. zebra* mitochondrial lineages.

Table 3.15. Polymorphic sites in 781 bp of mountain zebra of control region sequence.

```
                          10        20        30        40        50
                  ----|----|----|----|----|----|----|----|----|----|----
AUASBERG1         AAATCCCTACGAAGTGCCAGCCTGTATG-TTTTGCCACTAA--TCGCCTGATAA
AUASBERG2         AAATCCCTACGAAGTACCAGCCTGTATG-TTTTGCCACTAA--TCGCCTGATAA
AUASBERG3         A--CCCCTACGAGATATCGGCTCACGAA-TTTTGCCACTAA--TCGTCCGGCAA
CRADOCK1          AAATCCCTAAAAAGTATCGGCCCACACG-TTTCGCCACTAA--TCGCCTGGCAA
CRADOCK1          AAATCCCTAAAAAGTATCGGCCCACACG-TTTCGCCACTAA--TCGCCTGGCAA
DEHOOP1           AAATCCCTAAAAAGTATCGGCCCACACG-TTTCGCCACTAA--TCGCCTGGCAA
DEHOOP2           AAATCCCTACGAAGTATCGGCCTACATG-TCTCGCCACTAA--CCGCTTGGCAA
ERONGO            AAATCCCTACGAAGTACCAGCCTGTATG-TTTTGCCACTAA--TCGTCTGATAA
GAMKA             AAATCCCTACGAAGTATCGGCCTACATG-TCTCGTC-----TTCCGTCCGGCAG
GAMSBERG1         AAATCCCTACGAAGTATCGGCCCACATG-TCTCGCCACTAA--CCGCTTGGCAA
GAMSBERG2         A--TCTCCACGAAGTATCGGCCTACACAGTTTTACTACTAA--TTGCTCAGCAA
GAMSBERG3         A--TCCCTACGAAGTACCAGCCTGTATG-TTTTGCCACTAA--TCGCCTGATAA
KAMMANASSIE1      AAATCCCTACGAAGTATCGGCCTACATG-TCTCGCCACTAA--CCGCTTGGCAA
KAMMANASSIE2      AAATCCCTACGAAGTATCGGCCTACATG-TCTCGCCACTAA--CCGCTTGGCAA
KHOMAS HOCHLAND1  A--TCTCCACGAAGTATCGGCCTACACA-GTTTACTACTAA--TTGCTCAGCAA
KHOMAS HOCHLAND2  AAATCCCTACGAAGTACCAGCCTGTATG-TTTTGCCACTAA--TCGCCTGATAA
KHOMAS HOCHLAND3  AAATCCCTACGAAGTACCAGCCTGTATG-TTTTGCCACTAA--TCGCCTGATAA
KAMANJAB          AAATCCCTACGAGAAGTTGGCTCACACG-TCCTGCCACTAA--TCGCCTGGCAA
NAUKLUFT2         AAATCCCTACGAAGTATCGGCCCACATG-TCTCGCCACTAA--CCGCTTGGCAA
NAUKLUFT3         -AATCCCTGCAAAGTATCGATTCATACG-TTTTGCCACTAA--TCGCCTGACGA
NAUKLUFT1         AAATCCCTACGAAGTACCAGCCTGTATG-TTTTGCCACTAA--TCGCCTGATAA
OTJIVASANDU1      AAATCCCTACGAAGTATCGGCCCACATG-TCTCGCCACTAA--CCGCTTGGCAA
OTJIVASANDU2      AA-TTCTTGCGGAGTATCGATTTATACG-TTTTGCCACTAA--TCATCCGACAA
```

3.3.5.2 Mismatch distribution

The observed distribution of pair-wise differences between *E. zebra* haplotypes was not indicative of a sudden population expansion (expected curve). Instead, the observed distribution was multimodal and ragged, indicating a population in equilibrium.

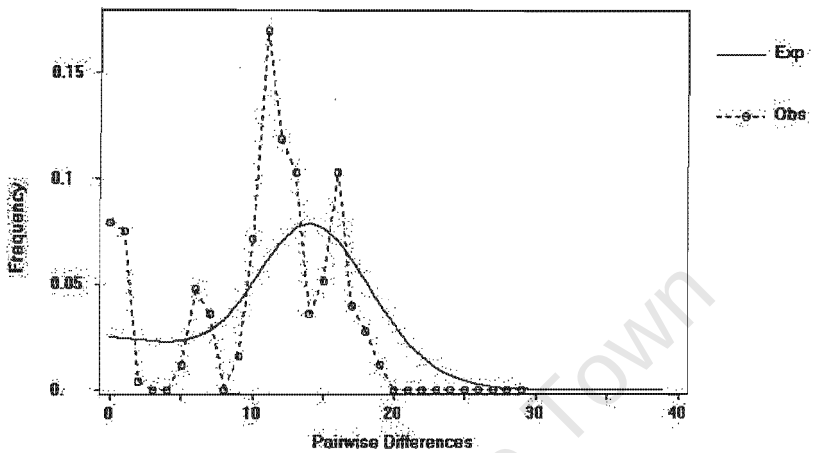

Figure 3.12. Mismatch distribution of pair-wise differences between mountain zebra mitochondrial haplotypes. Observed (Obs) pair-wise differences differ to that expected (Exp) if population size recently expanded.

3.3.5.3 Analysis of molecular variance (AMOVA)

The 23 mountain zebra sequences spanned 11 populations. An AMOVA tested the hypothesis that these populations were divided into three phylogeographic groups differentiated from each other by means of isolation by distance. Cape mountain zebras (n=7) sampled from Cradock, Kammanassie, Gamka and DeHoop were grouped against a Central Hartmann's group (n=13) consisting of Khomas Hochland, Gamsberg, Naukluft, Erongo and Auasberg and a Northern Hartmann's group (n=3) of Kamanjab and Otjivasandu. A moderate Φ_{ST} value of 0.1432 described the proportion of the variance in the data between populations relative to the proportion within populations. After 10 000 permutations, Φ_{ST} was not significant ($p > 0.05$), indicating that the probability of obtaining a value lower than the observed one was higher than can be expected due to chance. The control region therefore provides no support for phylogeographic partitioning of mitochondrial lineages.

3.3.5.4 Isolation by distance

The correlation between geographic distance between sampling locations and genetic distance between mitochondrial haplotypes, although weak (R^2 = 0.1409, Fig. 3.13), was significant ($p < 0.05$, Mantel test, after 10 000 permutations). This is consistent with the genetic structure observed with nuclear microsatellite data.

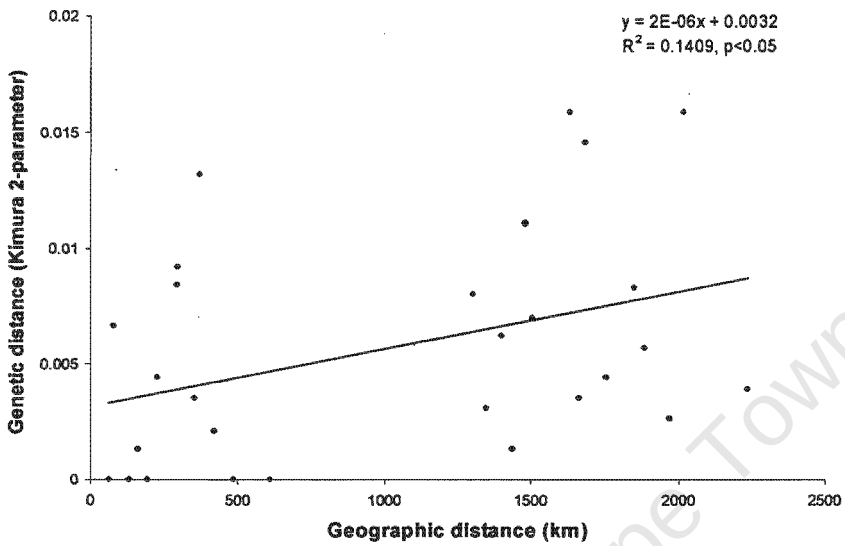

Figure 3.13. Isolation by distance among mountain zebra mitochondrial haplotypes.

3.3.5.5 Phylogenetics

There was no significant difference in log-likelihood scores between quartet puzzling trees with a molecular clock and trees without ($p > 0.05$, χ^2 test). Therefore the maximum likelihood (ML) tree calculated (Fig. 3.14) assumed a molecular clock which enabled the subsequent computation of coalescence time. The ML tree was thus scaled according to likelihood branch length which measures the average pair-wise distance between two taxa or groups of taxa. Both trees were rooted by an *E. caballus* outgroup taxon. Subjecting the data to both ML and maximum parsimony (MP, Fig. 3.15) criteria resulted in unresolved phylogenetic hypotheses, making it difficult to assign certainty to deep nodes. However, as some terminal nodes are supported by likelihood scores or bootstrap values, which are not compromised by the lack of resolution, certain conclusions may be drawn.

Almost all supported groupings are consistent with either method of tree construction, the only exception being that the Naukluft3/Otjivasandu2 and the Gamsberg2/Khomas1 clades are both unresolved under ML. The cladistic approach attains a higher level of

resolution than that of ML. This may seem surprising as this method makes use of only 29 parsimony informative synapomorphies in a data set with high haplotype diversity where homoplasy is expected to reduce the effectiveness of MP.

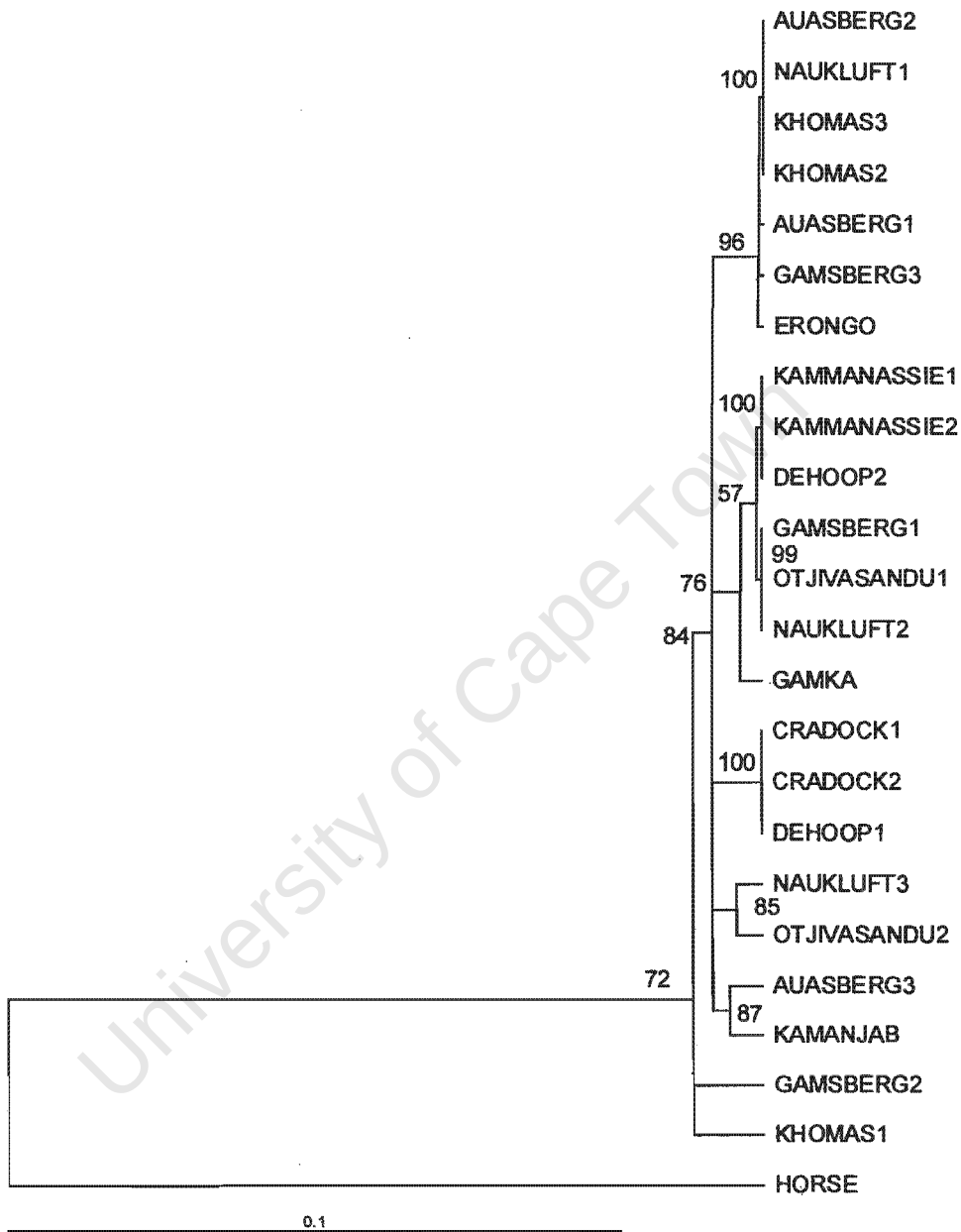

Figure 3.14. 50% majority rule consensus maximum likelihood phylogenetic hypothesis for 23 mountain zebra haplotypes. The horse (*E. caballus*) was chosen as the outgroup. Maximum likelihood branch lengths were calculated assuming a molecular clock. Quartet puzzling values are given for internal branches with greater than 50% support.

Mountain zebra mitochondrial lineages appear very divergent from the horse outgroup (Fig. 3.14), so confirming observations by Oakenfull *et al.* (2000). Consequently, neither tree resolved the observed pattern of isolation by distance (Fig. 3.13) as most haplotypes are very closely related. Maximum likelihood supported a clade of haplotypes from the Central region but the corresponding clade under MP contained a Northern region haplotype (Otjivasandu2). Two of the three Auasberg haplotypes fall into this clade but the third haplotype groups consistently with a Northern region haplotype from the Kamanjab population. The origin of the seeded Auasberg population therefore cannot be determined from control region sequences. The three Northern region haplotypes of Otjivasandu1 and 2 and Kamanjab are interspersed among three different groups, each with moderate support.

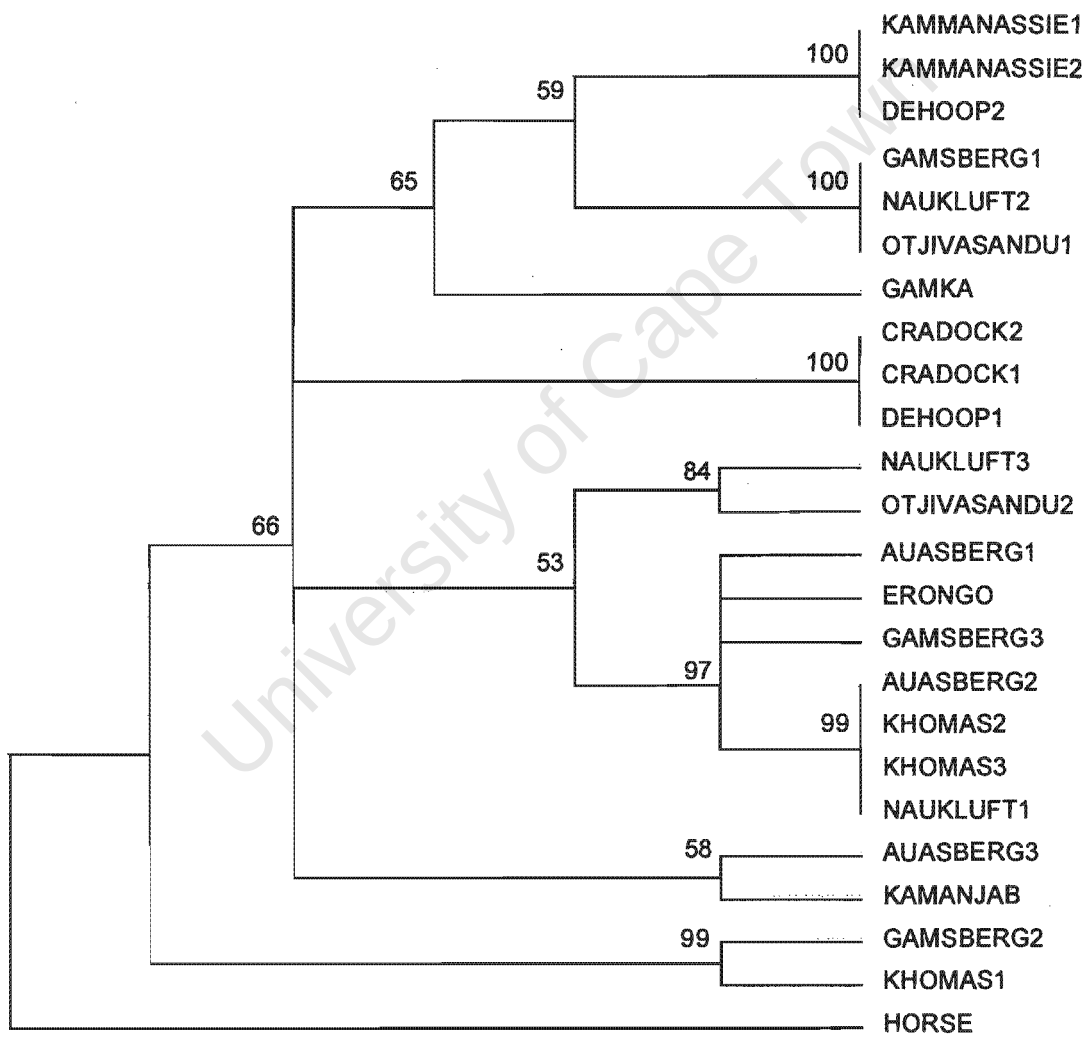

Figure 3.15. 50% majority rule consensus maximum parsimony phylogenetic hypothesis for 23 mountain zebra haplotypes. The horse (*E. caballus*) was chosen as the outgroup. Only nodes supported by more than 50% of 500 bootstrap replicates are shown.

Among what has hitherto been known as Cape mountain zebras, the three distinct haplotypes are characteristic of the three relict stock populations. Regardless of the lack of resolution of the data, both methods of tree reconstruction reject the hypothesis that Cape mountain zebras are monophyletic.

3.3.5.6 Coalescence

Genetic divergence between the horse and mountain zebras was corrected for rate heterogeneity using α = 0.12 (Oakenfull *et al.* 2000). The average nucleotide divergence, D_{xy} (Nei, 1987), between the horse outgroup and mountain zebra haplotypes was calculated using the Kimura 2-parameter model (Kimura, 1980) of nucleotide substitution as 0.11251 (± 0.02517) and by maximum likelihood analysis assuming the HKY model of nucleotide substitution as 0.12333 (± 0.04618). The fossil record for *E. zebra* is sparse. The only known *E. zebra* fossils were found in the south-western Cape Province of South Africa and date back to the late Pleistocene/early Holocene, 10 000 – 11 000 years ago (Hendey, 1974). Using the estimate of Oakenfull *et al.* (2000) for the emergence of the first horse-zebra ancestor at 2.3 Mya and applying $\mu = D_{xy}/2T$, nucleotide substitution rates of 2.45 x 10^{-8} substitutions/site/year (2.45 %/Myr) and (by ML analysis) 2.68 x 10^{-8} substitutions/site/year (2.68 %/Myr) were calculated for the data. Substituting the new μ values and using the two estimates of π within the mountain zebra data, 332 653 years (95% CI: 200 204 – 465 101 years) or (by ML) 362 313 years (95% CI: 217 537 – 507 089 years) were calculated as the time to coalescence of studied mountain zebra mitochondrial lineages.

3.4 DISCUSSION

3.4.1 Sampling effects

The most valuable results from these genetic data pertain to the genetic composition of, and relatedness amongst, the three Cape mountain zebra stock populations of Gamka, Kammanassie and Cradock. Sample size for Gamka and Kammanassie are small (n=9), yet the data suggest that these populations may contain up to two thirds of the genetic variation of the Cape mountain zebra metapopulation. Both populations decreased in number to no more than five individuals about 30 years ago. Comparisons with a seeded Hartmann's mountain zebra population (Auasberg) tend to suggest that Cape mountain zebra populations have been subjected to moderate drift since before each population was reduced to its lowest number. The historic literature (Woods, 1960) estimates this time to be over four generations (since about the turn of the 20th century) for Gamka and Kammanassie. It is therefore unlikely that the founders of the smaller Cape mountain zebra stocks contained much more genetic variation than that which is observed here. That each of these samples is in Hardy-Weinberg equilibrium for all tested loci also indicates that an adequate number of individuals was sampled. Furthermore, these genetic samples were collected by various means (see Section 2.2) over the course of at least 10 years. The possibility of obtaining a significant additional number of samples from these small and vulnerable populations is almost negligible. The present study is therefore likely to be the most comprehensive possible for defining the current genetic status of the Cape mountain zebra.

Sample size effects may have confounded results from the Naukluft (n=9) and Erongo (n=5) Hartmann's mountain zebra populations. Both these mountain ranges contain large numbers of mountain zebras and it is unlikely that these small sample sizes provide adequate estimates of population allele frequencies. This may be a factor influencing the lack of intraregional phylogeographic structuring in the Central region.

3.4.2 Statistics

3.4.2.1 Population differentiation

Fisher's exact test was found to be robust for the detection of population differentiation as it returned significant values for all but two pair-wise tests. This test may be oversensitive as significant differentiation was suggested between populations with low (<0.05) pair-wise fixation indices. The magnitude of the differentiation may be ascertained indirectly from p-values but owing to this suggested over-sensitivity, magnitude may only be deduced between a few of the least differentiated populations. It is suggested that the test be used initially to determine the presence or absence of population differentiation. Fixation indices and genetic distance may then be implemented to quantify the magnitude of the differentiation.

3.4.2.2 Genetic distances

Since the reliability of any distance measure is based on expectation and variance, it can be concluded from the data that D_C and $(\delta\mu)^2$ were the most appropriate measures of genetic distance at the intraspecific level. The pronounced effect of low within-population genetic diversity on D_S highlighted the unsuitability of this measure in tests of isolation by distance. All three distance statistics were affected by the small sample size of some Hartmann's mountain zebra populations. The high variance associated with $(\delta\mu)^2$ made it necessary that confidence limits always be used when estimating coalescence times.

3.4.3.3 Gene flow

Discrepancies between both estimates of gene flow make it difficult to draw realistic conclusions on the effective number of migrants between free ranging mountain zebra populations. In addition, the exacerbating effect of small sample size confounds any logical pattern of north-south gene flow among Namibian populations.

Wright (1931) showed that two populations would not diverge significantly if they exchange one or more migrants per generation. Applying this "one migrant per generation" rule reveals $N_e m_{RST} > 1$ in all comparisons of Hartmann's mountain zebra populations and $N_e m_{PA} > 1$ in 8/15 comparisons. As all natural populations which have not undergone any significant reduction in population size are assumed to be in

mutation-drift equilibrium, the data suggest that homogenizing gene flow still occurs between Hartmann's mountain zebra populations despite recent human encroachment and the erection of cattle fences.

Estimates of gene flow between the Cape and Hartmann's mountain zebra populations are likely to be underestimates of historic levels as the bottlenecked Cape mountain zebra metapopulation would be likely to have contained greater diversity in the past. Nevertheless, global N_em values for both estimates are less than one and therefore suggest that gene flow among mountain zebra populations has not been of a high enough magnitude to attenuate genetic divergence completely. The resulting pattern of isolation by distance at neutral microsatellite loci is thus a product of persisting genetic drift across the mountain zebra distributional range.

3.4.3.4 Population assignments and usefulness in determining structure

Assignment tests are usually employed in studies of introgression or hybridisation where an individual's multilocus genotype may be used to assess that individual's population of origin. In this study, the assignment test was used as a simple and indirect measure of population fidelity or structure. Although this interpretation of assignment tests has been criticised (Schneider *et al.* 2000), the fact that the results corroborate observations from other statistics implies that the information contained in multilocus likelihood products may be cautiously exploited in this fashion.

3.4.3.5 PCA

The multidimensional approach of PCA also provided support for conventional population statistics. PCA was able to describe the grossly differentiated Cape mountain zebra stocks as well as the moderate inter-regional structuring between Hartmann's mountain zebra populations. The statistically tested results obtained from this analysis therefore provide a concise graphical summary of the data.

3.4.3.6 Bottlenecking

Although Cape mountain zebra populations, with the exception of DeHoop, show population bottlenecking as defined by Garza and Williamson (2001), all the values obtained were higher than that reported for reduced or founder populations ($M < 0.7$). This may be explained when the number of loci used to compute M in each sample is

considered. Only polymorphic loci may be used to compute *M* as monomorphic loci have a spread equal to the number of alleles present, thereby resulting in an inappropriate *M* value of 1. The test is therefore biasing the results towards higher values by removing the very loci that are most affected by population bottlenecking. Owing to a rapid loss of allelic diversity in each stock, between only six (GDNR) and 12 (DeHoop) of the 15 possible loci were used for any of the Cape mountain zebra populations despite the entire metapopulation being polymorphic for 14 loci (Table 3.2B). The method hence rejected much of the very data that it requires to demonstrate the gross level of inbreeding that follows a severe genetic bottleneck. This is a serious flaw in this otherwise promising method and it is suggested that this problem be addressed by its developers. Even the allocation of an arbitrary *M* value of 0.5 to each monomorphic locus significantly decreases the overall *M* in all Cape mountain zebra populations (*M* would then = 0.531 - 0.590). By simulating population bottlenecks and plotting average *M* values for decreasing allele size, one may be able to extrapolate the "ideal" *M* value for a monomorphic locus.

3.4.3.7 Nuclear and mitochondrial coalescence estimates

The highest nuclear divergence estimate between any two mountain zebra populations was 242 792 years (upper confidence limit). Although this falls within the lower confidence limits of both control region estimates (200 204 and 217 537 years), the nuclear estimate is biased by the small sample size of the Erongo population and by the marked differentiation between individual Cape mountain zebra populations (Gamka in this case) and Hartmann's mountain zebra populations. The highest nuclear coalescence time estimate between the Cape mountain zebra metapopulation and Hartmann's mountain zebra populations (86 967 years) more realistically illustrates the ability of $(\delta\mu)^2$ to detect historic signal in microsatellite data. Given that the initial assumption of an average mutation rate at microsatellite loci of 2.05×10^{-4} (Lehmann *et al.* 1998; Rooney *et al.* 1999) is relatively conservative, a substantial departure from linearity of the $(\delta\mu)^2$ statistic is evident. The nuclear structure of free ranging mountain zebra populations is characterised by high gene flow and relatively low genetic drift. While drift manifests itself by producing more divergent populations with increasing geographic distance, gene flow tends to have a homogenising effect, keeping the genetic content of populations relatively similar. While non-recombining haplotypes reveal a linear maternal evolutionary history, recombining microsatellite loci lose linearity relatively quickly as back mutations mask the extent of mutational change.

Owing to its development with regard to the SMM, $(\delta\mu)^2$ is assumed to remain linear for longer periods than other distance statistics. At the same time, however, the discrepancy between the nuclear and mitochondrial coalescence estimates suggests that $(\delta\mu)^2$ plateaus sooner than expected.

3.4.3 Mitochondrial population structuring

The high incidence of between-population haplotype sharing resulted in a weak isolation by distance correlation coefficient. Despite a greater within-population variance and consequent lack of significance of Φ_{ST} when the sequence data are subjected to AMOVA (Section 3.3.5.3) and a lack of complete resolution in both intraspecific phylogenies, significant IBD was detected in the *E. zebra* mitochondrial data. This indicates a clinal distribution of genetic heterogeneity along the *E. zebra* distribution. This in turn suggests that the magnitude of female gene flow is compatible with the mean of that of both sexes.

Intraspecific haplotype sharing has been documented by Oakenfull *et al.* (2000) in hemione stenoids, but in the present study, all three Cape mountain zebra haplotypes occur exclusively within the subspecies *zebra*. Nevertheless, the explanation tabled by Oakenfull *et al.* (2000) that a recent divergence has resulted in two genetically similar forms with too little distinctiveness to classify them as separate subspecies may adequately explain the lack of monophyly observed in *E. zebra* subspecific haplotypes.

3.4.4. Nuclear population structuring

3.4.3.1 Northern (Hartmann's) mountain zebra

Moderate to low population structuring was observed within the Namibian mountain zebra metapopulation over the 15 studied microsatellite loci. Fixation indices and genetic distances loosely defined two sub-groups, the Northern populations of Otjivasandu and Kamanjab and the Mid-Central populations of Khomas-Hochland and Gamsberg. These groupings were qualitatively supported by likelihood assignment tests and by principal component analysis. The placement of the North-Central and South-Central populations of Erongo and Naukluft is not clear and more samples are required from both these populations. The three distance statistics show significant support for the hypothesis of isolation by distance across the mountain zebra range. The strongest correlation with geographic distance was observed for D_C, and given that

this statistic is least affected by within-population genetic diversity, it can be assumed that the differentiation among populations in Namibia is caused by differential drift owing to isolation by distance. It may also be proposed that this mechanism describes the historic nuclear genetic structure of mountain zebra populations across their entire range.

That livestock fences were hindering the inter-regional exchange of genetic material has long been a concern for the Namibian Ministry of the Environment and Tourism (MET). In the dryer months from August to October, the most accessible permanent water on the Namib Escarpment and the Central Plateau is found at installed watering points for sheep and cattle. These areas are often fenced off for exclusive use by livestock. Due to their dependence on water, mountain zebras have been known to break fences in order to reach these sources. Many farmers in the Khomas Hochland now leave gates open to watering points in order to prevent this destruction. Their ability to destroy traditional livestock fencing renders these barriers unlikely hindrances to the movement of mountain zebras in Namibia.

The increasing popularity of game farming may mean that sturdier game fences may succeed in reducing gene flow causing local and perhaps regional insularity and differentiation. The results here suggest that this may have already taken place in the Erongo Mountains as this population shows signs of the heterozygosity excess characteristic of a recent reduction in effective population size (Luikart *et al.* 1998).

Smaller unstudied populations in the southern extremities of Namibia and in the Richtersveld have been reduced to small numbers by human encroachment from as early as the late 1800s and by uncontrolled poaching. These populations have probably undergone accelerated drift due to this decrease in population size and are expected to show higher levels of differentiation relative to Northern and Central regions than could be expected due to a north-south clinal pattern of isolation by distance alone.

Apart from these populations in southern Namibia and apparently those in the Erongo Mountains, humans appear to have had little impact on the nuclear genetic composition of Hartmann's mountain zebra populations. This, as in most developing countries with a growing population, is likely to change in the future. However, given the results from the present study, the MET still has a number of management options to curb the

future loss of genetic variation. Game fencing needs to be controlled and corridors set up to aid the movement of those animals outside game farms. More watering points need to be installed in dry river valleys to encourage mountain zebras away from farmland where they may be hunted. Stricter anti-poaching measures also need to be enforced in remote areas.

3.4.3.2 Southern (Cape) mountain zebra

The data show that the Cape mountain zebra metapopulation still retains a proportion of its historic genetic variation. Component stock populations of this South African metapopulation, however, show a marked decrease in genetic variation compared to the free-ranging populations in Namibia. The Cape mountain zebra populations were also found to be highly structured and differentiated relative to each other. Unlike the moderate genetic structuring between the entire Hartmann's and Cape mountain zebra metapopulations, the gross levels of differentiation between the Cape mountain zebra stocks cannot be explained by isolation by distance nor by any naturally occurring founder or vicariance event. The extant Cape mountain zebra stock populations were reduced to very small numbers and isolated from each other since as early as the late 1800s. Results show that this prolonged population bottleneck promoted accelerated drift, thereby explaining the marked differentiation of these populations from each other.

The current genetic status of the Cape mountain zebra is thus a product of almost complete population decimation. The proclamation of the MZNP in 1937 served to protect only a fraction of the extant gene pool while other potentially variable populations in the Outeniqua and Kouga Mountains were left to become extinct. A resurgence of the need to conserve hitherto unprotected populations of Cape mountain zebra finally surfaced in the late 1960s, and resulted in the survival of the relict populations of Gamka and Kammanassie.

Although the results here illustrate the desperate genetic situation of individual Cape mountain zebra stock populations, they also show that historic levels of genetic diversity have not been altogether lost from the Cape mountain zebra metapopulation. Maintaining this diversity is critical to the survival of the subspecies as inbreeding depression may already have set in in these remnant populations and further inbreeding may lead to an even greater decrease in fitness. Already, the appearance

and extensive spread of a form of equine sarcoid in the Gamka, BNP and GDNR populations is consistent with inbreeding and is a cause for concern. Even though the causal link between BPV and the BNP/GDNR sarcoid outbreak has not yet been demonstrated, the causative virus is suspected as having been contracted from cattle which graze on the pastures bordering both reserves. Given that the genetic predisposition for BPV and sarcoid oncogenesis has been demonstrated more than once (Lazary *et al.* 1985; Meredith *et al.* 1986; Brostrom *et al.* 1988; Lazary *et al.* 1994) and that the presence of sarcoid lesions may impact on individual fitness, pro-active management of Cape mountain zebra populations has never been more essential.

Until now the management of the Gamka and Kammanassie populations has followed a "hands off" approach allowing the populations to recover naturally with minimal interference. Unfortunately, neither Gamka Mountain Nature Reserve nor Kammanassie Nature Reserve constitutes a natural, functioning system. For example, most of the zebras on the Kammanassie Mountains occur on the less rugged western end close to the town of Dysselsdorp. Two mares, in peak condition roam the steep eastern slopes, isolated from the rest of the population. The relocation of these mares to the populous western sector or the moving of a young stallion to the eastern sector has long been advocated by local officials. By contrast, the opposite extreme has occurred at the MZNP where SANP discontinued the use of an official studbook and began large-scale movement of mountain zebras to satellite populations. Results show that this practice, while rewarding in aiding the rapid expansion of the Cradock-derived metapopulation, also subjected 25 usually small satellite populations to a secondary bottleneck. This not only kept genetic diversity low but also could have removed variation that was probably present prior to the translocations. The presence of unique alleles at DeHoop confirms that the MZNP and possibly Kammanassie were more genetically diverse in the 1970s than either is at present. This diversity has subsequently been lost from these respective stocks owing to the rapid drift consequent to small population size.

If the *status quo* is maintained and each stock remains isolated, drift will continue to remove genetic variation from each population until those few microsatellite loci with variable alleles also become fixed. The populations at Kammanassie and Gamka are not large enough to attenuate this accelerated drift. The MZNP and its derivatives make up a population large enough to curb the effect of drift or at least to reduce the rate at which alleles are lost. However, even the larger Cradock-derived populations

such as KNP, KNR and MZNP will also drift toward fixation, albeit at a rate slower than that of the smaller Gamka and Kammanassie, unless their population numbers are significantly increased (to values given in Table 3.13). It is therefore clear that the genetic components of each Cape mountain zebra stock have to be combined in order to increase the diversity (and hence, fitness) in a combined population.

Indicators of inbreeding depression

The Kammanassie:Cradock mixture of the DeHoop population provides a unique window that allows us a temporal snap-shot, 30 years into the future, of the genetic consequences of mixing stock populations today. DeHoop is the most genetically diverse of the Cape mountain zebra populations and has shown steady but not rapid population growth since its founding (Fig. 3.5, Novellie *et al.* 2002). As other less diverse Cradock-derived satellite populations have shown more substantial population increases, population growth is most likely to be as much a function of habitat quality as a decrease in inbreeding depression, and it may be difficult to differentiate between the two.

Equine sarcoid disease may also not appear to be a completely unreliable general measure of inbreeding depression as it is exhibited in not only the most genetically depauperate Cape mountain zebra population (GDNR), but also in the moderately diverse BNP population. It may be that only some populations lose the critical "defensive" alleles that allow sarcoid oncogenesis. However, the contraction of BPV appears to be dependent on the close proximity of cattle. The present lack of neutral genetic variation in Cape mountain zebra populations may translate to levels of MHC variation that are low enough to promote sarcoid growth in any population that is exposed to cattle. There is also evidence to suggest that equine sarcoid is readily transmitted between animals in close contact (Reid *et al.* 1994) and as a result, steps should be taken to remove cattle from areas bordering Cape mountain zebra reserves.

3.4.5 Cape mountain zebra management strategy

Seeding of new satellite populations from the MZNP will only add to the present network of genetically depauperate populations and allow accelerated genetic drift to target the small founder populations. As exemplified by the population at DeHoop, effective management of Cape mountain zebra genetic diversity can only be achieved if stocks are combined efficiently. In the formation of such management strategies,

genetic results should be used in conjunction with existing ecological and demographic information. Thus, selected animals, preferably males from the two smaller stocks of Kammanassie and Gamka, could be paired with females, preferably more than one, on new reserves. It would be less valuable to introduce males from Kammanassie or Gamka to existing Cradock-derived populations since mountain zebra social dynamics indicate that these males are very unlikely to become herd stallions and sire an optimum number of foals (Penzhorn, 1984b; Lloyd and Rasa, 1989).

Selection of new reserves for seeding mixed herds will depend on factors such as an abundance of adequate grazing and water. Initial reserve size and terrain may also be critical. If possible, seeding should occur in a small reserve of gradually undulating mountainous slopes similar to Commando Drift Nature Reserve where unprecedented Cape mountain zebra population growth has taken place in the last 10 years (Novellie *et al.* 2002). Situations such as that of the Kammanassie Nature Reserve where some animals are separated from members of the opposite sex by vast distances and rugged terrain should be avoided. Conversely, terrain must be rugged enough to inhibit excessive hoof growth. Should population growth increase rapidly, the expansion of the reserve would have to be considered. As demonstrated by the present study, the expansion of an existing population is more beneficial than translocating a subset of that population to another reserve, thereby subjecting the founding animals to a secondary bottleneck.

Given the present numbers at Kammanassie and Gamka, it will be impossible to seed a large herd. Intermittent augmentation of populations is therefore desirable to prevent the inbreeding of small groups. It may take some months for females to accept a foreign stallion but once this takes place other, more common Cradock-derived males may be introduced into the population. In this way, the natural socio-dynamic process would be simulated without the risk of the new stallions contributing to the gene pool. These new reserves will need to keep an updated studbook detailing all aspects of population dynamics and demographics.

Cattle must be removed from areas bordering both newly seeded and presently existing reserves in order to remove the threat of BPV induced sarcoid diseases and its possible effect on mountain zebra fitness. The low success of surgical excision, the large excision area required and the risk of autoinoculation (Brostrom, 1995) make such a method of treatment undesirable. Non-specific immunostimulants have been

successful in aiding complete or at least partial regression of equine sarcoid in horses (Studer *et al.* 1997) and it is therefore proposed here that this form of treatment be administered to those mountain zebras already infected with sarcoid tumours.

The implementation of this management strategy may prove expensive. However, as SANP invariably sets aside an annual budget for the translocation of Cradock-derived Cape mountain zebras, the focus of these translocations need only shift to the movement of females from the larger Cradock-derived populations, particularly the MZNP and KNR and males from the two smaller stocks. Over the last few years, a degree of collaboration has materialised between SANP and provincial nature conservation authorities of the Western and Eastern Cape Provinces. This will be essential if the management of Cape mountain zebra is to be successful. Conversely, the involvement of the private sector would also be beneficial as game farmers often have to cover the costs of translocation. Before this involvement can happen, the relevant provincial authorities must first ascertain whether the prospective private reserve meets the criteria set out above.

It is evident that, although the Cape mountain zebra metapopulation has retained much historic genetic variation, this level is still not as high as that observed in Hartmann's mountain zebra populations. If the variation within the Cape mountain zebra metapopulation cannot be maintained as suggested above, the results from this study provide no scientific reason (see Figs 3.14 and 3.15) why the Cape mountain zebra gene pool may not be augmented with genetically variable Hartmann's mountain zebras.

3.4.6 Conclusions

Equus zebra once maintained a continuous distribution from south-west Angola south to the foothills of the Drakensberg. Throughout its range, mitochondrial haplotypic variation and nuclear microsatellites suggest a significant phylogeographic pattern of isolation by distance. Evolutionarily, extant lineages were found to be of recent origin, with a high diversity of closely related haplotypes pointing to a recent coalescence approximately 200 204 – 507 089 years ago. The species was divided into Cape and Hartmann's subspecies on the basis of phenotypic variation. The plasticity of these phenotypic variables has been outlined (Rau, 2002) and confirmed by the absence of support by both mitochondrial and nuclear molecular data for subspecific delineation.

Nevertheless, the habitats populated by the northern and southern forms, while containing all the prerequisites necessary for mountain zebra habitation, are significantly different enough to have possibly conferred slight selective differences to their respective forms which are too recent to be manifested in neutral DNA markers. It is therefore proposed that, while there is no DNA evidence to support the existence of *zebra* and *hartmannae* subspecies, if the genetic diversity in the South African metapopulation is adequately maintained, there is no obvious advantage to managing the entire species as a whole.

Recent human mediated decimation of populations in the south of the *E. zebra* range resulted in artificially high levels of nuclear microsatellite variation between individual southern populations and all others. The northern or Hartmann's mountain zebra appears not to have suffered the population crashes of the southern or Cape mountain zebra. The northern populations are still largely free ranging and show high levels of within-population genetic variation and low levels of between-population variation. There is genetic evidence that this situation is already changing in some Namibian populations in response to an increasing human population and perhaps to the rise in popularity of game farming and consequently the construction of game fences. Namibian authorities can no longer expect to oversee large, genetically diverse populations of mountain zebra in the future without significant management and planning.

In South Africa, this molecular genetic study has shown that Cape mountain zebra populations, while singularly depauperate and inbred, collectively contain close to historic levels of allelic diversity. The results have demonstrated that levels of diversity have dropped even in the last 30 years, owing to genetic drift. If current levels of diversity are to be maintained without further loss, genetic components of individual stocks should be combined to increase heterozygosity and alleviate the latent effects of inbreeding. It is imperative that these measures be implemented as soon as possible.

Chapter 4: Population genetic structuring in southern African plains zebra (*Equus quagga*)

4.1 INTRODUCTION

Equus quagga, the plains zebra, is the most common and successful of all wild equids, inhabiting the vast tracts of African savanna from Equatoria Province in the Sudan to the Karoo in South Africa. Described as *Equus burchelli* by Gray in 1824, its classification warrants re-appraisal after recent molecular evidence showed that it was indistinguishable from *Equus quagga*, the extinct Cape quagga (Higuchi *et al.* 1987). As the quagga was first described by Boddaert in 1785, the name *E. quagga* should take precedence over *E. burchelli*.

Figure 4.1. Damara plains zebra (*Equus quagga antiquorum*) in Chobe National Park, Botswana. Distinguishing features depicted here include the long mane terminating in front of the ears, dark muzzle in adults, diminished striping on the limbs, shadow stripes within white regions on the rump and hindquarters and a paucity of stripes on the lower limbs.

Equus quagga is a complex species assigned from four (Cabrera, 1936; Dorst and Dandelot, 1970) to six (Groves, 1974) and seven (Antonius, 1951; Ansell, 1974; Kingdon, 1997) subspecies based on variation in pelage striping and colour. The nature of pelage variation in the *E. quagga* complex across its large distribution range appears clinal along a roughly north-south axis (Dorst and Dandelot, 1970; Rau, 1978).

4.1.1 Distribution and distinguishing features

Equus quagga is a large ungulate with a mean mass of 320kg, standing approximately 1.36m at the shoulder (Skinner and Smithers, 1990). In southern Africa, where they occur sympatrically with mountain zebras, *E. quagga* is distinguished by its larger size, dark muzzle, a mane that terminates in front of the ears and the absence of both a dewlap and the grid-iron pattern on the croup. At the other end of their distribution, in south-western Ethiopia, plains zebras may be found in sympatry with the much larger (Kingdon, 1997) and narrowly-striped Grevy's zebra (*E. grevyi*).

Intraspecifically, *E. quagga* varied from a broad and distinctly striped zebroid (*E. q. boehmi*) in the north of its range to a semi-striped equid (*E. q. quagga*, extinct) where the stripes on its back and hindquarters merged to produce a dull brown colour. Between these extremes, a number of potential subspecific patterns emerge (see Fig 4.2, page 108), probably owing to local drift as a result of isolation by distance. This has led to clinally distributed pelage and skull variation (Rau, 1978). Compounding this, the high individual variation in striping pattern, as demonstrated by Smuts (1974), was not known to early explorers and scientists. Thus, by the early 1900s, up to 64 different forms were described (Cabrera, 1936), each on the basis of a few skins. Numerous attempts have been made at ordering these various forms into a logical subspecific framework (Cabrera, 1936; Antonius, 1951; Dorst and Dandelot, 1970; Ansell, 1974; Groves, 1974; Kingdon, 1997) but all attempts have been confounded by the widespread intergradation between neighbouring subspecies due to a lack of significant barriers to gene flow historically.

Generally, the entire distribution of *E. quagga* may be reduced to three broad groups or foci (Kingdon, 1997), occurring in East, Central and southern Africa. At a deeper level, the loose guideline summarising presently conventional subspecific forms is given below and described in Figure 4.2. The East African 'group' is comprised of a single

subspecies, Boehm's zebra, sometimes referred to as Grant's zebra. *Equus quagga boehmi* ranges throughout the grasslands of East Africa, with the southern end of its distribution bound approximately by the Rovuma River and Lake Tanganyika. In the Central region, the Upper Zambezi zebra (*E. q. zambeziensis*) occurs south-west of this and west of the Luangwa River's Muchinga Escarpment to include the Barotese floodplains and eastern Angola. The other Central form, Crawshay's zebra (*E. q. crawshayi*), ranges from the Luangwa Valley to the Indian Ocean, north of the Zambezi River. Chapman's zebra (*E. q. chapmani*) is found from south of the Zambezi River to the Limpopo River but not in the Kalahari. This subspecies is also known as Selous' zebra (*E. q. selousi*). The rest of the southern African sub-region from South Africa and north-west across the Central Kalahari to northern Botswana, northern Namibia and Benguela is inhabited by *E. q. antiquorum*, the Damara zebra. The Chapman's and Damara zebras form what will be referred to as the southern African group. The southern African sub-region was formerly inhabited by two additional extinct subspecies, Burchell's zebra (*E. q. burchelli*) of the Northern Cape and Free State Provinces and the true quagga (*E. q. quagga*) of the Karoo and southern Free State Province.

In the northern *E. q. boehmi*, striping is broad and distinct and all animals are striped down to the hooves. Most of the variation in plains zebra pelage occurs outside of this group. The Upper Zambezi zebra is similar to Boehm's zebra but differs in the length of the brain case (Groves, 1974) and the infrequent occurrence of grey shadow stripes between black stripes. In Crawshay's zebra, stripes are narrower and more numerous and shadow stripes therefore, almost never occur. The mean shoulder height of zebras of the Northern and Central groups tends to be 6 - 8% less than either of the southern African forms (Smuts, 1974). The white spaces are wider within the Southern group allowing for the more frequent occurrence of shadow stripes. The two southern African subspecific forms differ in that Chapman's zebra are striped down to the hooves whereas Damara zebras are often incompletely striped below the elbow and stifle (Dorst and Dandelot, 1970; Skinner and Smithers, 1990). The Damara zebra shows the highest variety of striping morphs. In KwaZulu, South Africa, up to 15% of all *E. q. antiquorum* exhibit the extinct *E. q. burchelli* phenotype (Rau, 1978) of significant disruption and sometimes complete reduction of hindquarter stripes. At the other end of the Damara zebra range, there is an extreme widening of shadow stripes in Namibia and in some cases these may occur as far up as the neck and forequarters. It is

thought that this variation was the result of continuous gene flow between Damara zebras and both the Burchell's zebra and the true quagga (*E. q. quagga*) where striping of limbs and hindquarters was almost completely absent (see Fig. 4.2).

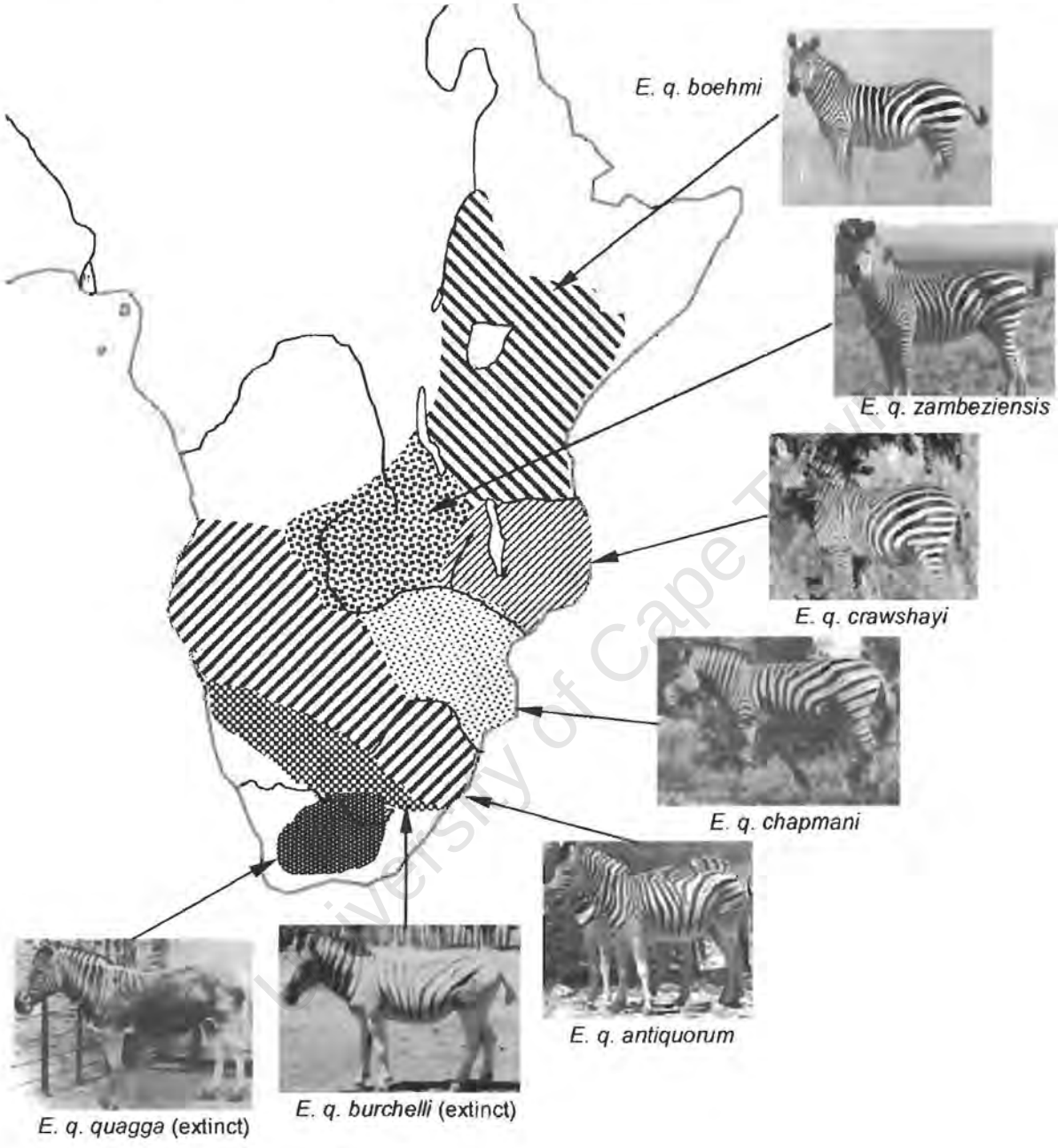

Figure 4.2. Subspecific distributions of plains zebra (*Equus quagga*) in Sub-Saharan Africa. As clinal and intrapopulation variation in stripe pattern is considerable, the above should act only as a rough guide.

In an increasingly fragmented habitat, the recognition of subspecific forms is helpful for conservation. However, to be truly useful, this classification must recognise areas of intergradation where two or more subspecific phenotypes may be breeding within the same local population.

4.1.2 Ecology

The plains zebra is predominantly a grazer but has been known to browse on occasion (Ansell, 1960; Smuts, 1972). Consequently they are partial to open or slightly wooded grasslands but their preference for different grass species varies with specific habitat. They are a water-dependent species and rarely range more that 10 -12 km from it. Although gregarious, *E. quagga* is not territorial. Instead they forage over seasonally plastic and overlapping home ranges which vary in size from 111 km^2 to 262 km^2 depending on the quality of habitat (Smuts, 1972). Plains zebras undertake seasonal migratory movements between summer and winter feeding grounds. In the Serengeti plains, the distances covered by these migrations may exceed 200 km (Grzimek and Grzimek, 1960). Plains zebras are often found in association with wildebeest (*Connocaetes taurinus*) and other plains grazers, presumably in order to avoid or dilute the probability of being selected for predation by lion (*Panthera leo*). Spotted hyaenas (*Crocuta crocuta*) are known to prey on foals (Smuts, 1974). In Etosha National Park, Namibia, plains zebra frequently associate with mountain zebra at watering points. Although respective herds do not mix, there have been reports of suspected hybrids from this area (R. Rau pers. comm.)

4.1.3 Social organisation

Plains zebras are organised into family groups, several of which may aggregate to form a herd. A family consists of a stallion and one or more mares and their foals. Colts are expelled from the family group after two to three years and form bachelor groups. In herd situations, bachelor groups are restricted to peripheral positions and rarely integrate with family groups. New family groups are formed when bachelors liase with young females. Most family groups are led by stallions 4.5 – 12 years old; however, in areas of high zebra density, stallions may not lead groups until about five or six years old (Skinner and Smithers, 1990). When threatened, the dominant stallion takes up a rear position in order to defend his group. The size of family groups is correlated with habitat condition and level of predation (Smuts, 1974). Family group size varies from an average of 7.7 individuals under optimal conditions (Turner and Watson, 1965) to four, under the influence of either habitat condition or predation (Smuts, 1974). Mares

are able to reproduce at three years old. Longevity is very similar to that recorded for mountain zebras (Jones, 1993). Generation time for both species was therefore identical (16.5 years). Foals are born throughout the year but Smuts (1974) provided evidence of a rainfall-induced peak in reproduction at the beginning of summer. Gestation varies from 360 to 390 days. Mortality is high in the first year of life, the primary influencing factor being predation (Smuts, 1974).

4.1.4 Population history

4.1.4.1 General

Until recently, plains zebra were able to move freely throughout their distributional range. With the erection of veterinary fences to prevent the spread of nagana and foot and mouth disease together with the increase in human population density, most plains zebra populations have become localised and concentrated in protected wilderness areas. In much of Africa, conservation areas are generally not fenced off from surrounding districts and wildlife is able to range freely within these extended areas. However, in South Africa, fenced parks and reserves further restrict the ability of plains zebras to undertake necessary seasonal migrations.

Table 4.1. *Equus quagga* numbers by country and subspecies. Adapted from Hack *et al.* (2002).

Country	Subspecies	Number (number in subspecies)	
Sudan	*E. q. boehmi*	(493 665)	33 050
Ethiopia	*E. q. boehmi*		7470
Somalia	*E. q. boehmi*		1000?
Kenya	*E. q. boehmi*		152 490
Uganda	*E. q. boehmi*		3137
Rwanda	*E. q. boehmi*		3048
Tanzania	*E. q. boehmi*		296 508
DRC	*E. q. zambeziensis*	(19 219)	1000?
Malawi	*E. q. crawshayi*	(23 020)	670
Zambia	*E. q. crawshayi/ E. q. zambeziensis*		39 469
Mozambique	*E. q. crawshayi/ E. q. chapmani*		1165
Zimbabwe	*E. q. chapmani*	(20 294)	20 135
Botswana	*E. q. antiquorum/ E. q. chapmani*		34 294
Namibia	*E. q. antiquorum*	(103 976)	13 090
South Africa	*E. q. antiquorum*		55 686
Swaziland	*E. q. antiquorum*		1000
TOTAL	***E. quagga***	**(660 174)**	**~663 212**

In the late 1800s, the rinderpest viral pandemic spread through Africa, decimating populations of buffalo and wildebeest while destroying thousands of cattle. As cloven-hoofed game were seen as the carriers of the rinderpest virus, large scale extermination of these animals was carried out. Although equids neither carried nor were affected by the virus, large numbers of zebra were culled. However, the bottleneck endured by the resulting fragmented groups was short-lived and populations rapidly recovered. Extant population census numbers across the entire *E. quagga* range are shown in Table 4.1.

4.1.4.2 Southern African sub-region

In southern Africa further population reduction occurred in the first half of the 20th century as a result of campaigns to eradicate nagana, a debilitating trypanosome-induced disease of cattle, vectored by the tsetse fly (*Glossina* spp.). As unaffected hosts of the trypanosome, wild game was seen as a reservoir from which constant re-infection of domestic cattle herds was inevitable. As agriculture was the primary means of income at that period, the campaigns focussed on the destruction of game species in wide areas of human habitation. In Southern Rhodesia (now Zimbabwe) alone, 659 334 game animals were slaughtered between 1919 and 1961 (Mann, 1990). However, relatively large numbers of plains zebra remained in the Zambezi Valley and populations recovered quickly. In Zululand, now known as KwaZulu, an estimated 15 000 zebra were destroyed in 1929 and 1930 and a further 1700 were culled between 1942 and 1950 with the aim of complete extermination (Mentis, 1970). The few remaining individuals founded the nucleus of the population that now inhabits the Hluhluwe-Umfolozi Park in northern KwaZulu. This Hluhluwe-Umfolozi population recovered once again and the population is now estimated to be about 3000 (Hack *et al.* 2002).

Usually a lack of gene flow across the plains zebra range would lead to eventual population localisation and differentiation due to genetic drift. This effect would be compounded by population reductions caused by demographic bottlenecks. If populations remain small after a bottleneck, the effects of drift on genetic diversity would be readily apparent. However, the plains zebra is one of the few large African ungulate species that is still plentiful, with a metapopulation size in excess of 660 000 (Hack *et al.* 2002). Despite localisation and population reductions, large local populations of *E. quagga* are common (Table 4.1). These large local populations will

slow down the effect of genetic drift. Therefore recent, human-mediated population pressures such as population fragmentation and the erection of veterinary fences may not result in significant changes in allele frequencies when population sizes are large.

The demographic bottlenecks sustained by the Hluhluwe-Umfolozi population may have reduced the population to a number small enough for the detection of a perceptible loss in genetic diversity. A proportion (15%) of the Hluhluwe-Umfolozi zebras exhibits reduced striping on the hindquarters that produces a phenotype similar to that of the extinct Burchell's zebra (Rau, 1978). Selection for reduced striping during a quagga breeding programme in Pretoria, South Africa, produced a collateral decrease in sperm count (Bowland *et al.* 2001), suggesting that this characteristic in wild populations may be indicative of a loss in diversity and a reduction of fitness. In South Africa, the rapid increase in game ranching for the purposes of hunting and ecotourism has resulted in a high proportion of plains zebras being confined to small founder populations, isolated by perimeter game fences. Most of the small private reserves on the South African Lowveld and in the erstwhile Transvaal Province have been seeded with animals from the Kruger National Park. As the only remaining wild population of plains zebra in KwaZulu, the Hluhluwe-Umfolozi population has been the source of no fewer than 12 such zebra populations. That two-thirds of these reserves were seeded with less than 20 individuals (P. Taylor, pers. comm.) would have subjected those founders to a secondary bottleneck, thus further accelerating drift.

Evidence of inbreeding depression has already been observed in some of these seeded populations founded with Hluhluwe-Umfolozi stock. A high proportion of stillbirths was reported for Vernon Crookes Nature Reserve, seeded with 19 animals; striping pattern similarity in Enseleni Game Reserve, seeded with six animals; and small size, reduced striping on the hindquarters and high mortality in Harold Johnson Nature Reserve, seeded with eight zebras. Knowledge of the genetic status of small isolated seeded populations is necessary in order to maintain diversity and lessen the effects of inbreeding depression. Bowland *et al.* (2001) carried out a PCR-RAPD, allozyme and PVA analyses on a comprehensive sample from Hluhluwe-Umfolozi and three populations derived from it. The analysis showed a loss of genetic variation in seeded populations that was negatively correlated with population size. An intensive metapopulation conservation strategy was advocated, with the frequency of recruitment dependent upon population size and genetic diversity.

4.1.5 Phylogeography

A number of mammal species with pan-African distributions appear to have a disruption in gene flow between northern and southern African populations (Girman *et al.* 1993; Matthee and Robinson, 1997, 1999; Kingswood, 1998). In a survey of equid control region and 12S rRNA mitochondrial DNA, Oakenfull *et al.* (2000) were unable to ascertain whether this was also true for plains zebras as East African and southern African groups were not monophyletic and geographically intermediate populations were not sampled. However, appreciable haplotypic population structuring has been observed between cytochrome oxidase gene sequences of Kenyan and Botswanan *E. quagga* (N. Georgiadis, pers. comm.). Due to contrasting results and wide sampling gaps, the phylogeography of *E. quagga* mitochondrial haplotypes across its distributional range is thus not yet fully understood.

In southern Africa, plains zebras exhibit the highest observed phenotypic diversity, with as many as four subspecies described. Only two of these, the Damara zebra (*E. q. antiquorum*) and Chapman's zebra (*E. q. chapmani*), are extant. In the absence of any barriers to gene flow between the distributional ranges of the two subspecies, there should be wide areas of intergradation, such as in western Zimbabwe and north-eastern Botswana where individuals within a family group may resemble representatives of either subspecies.

4.1.6 Aims

As shown in the previous Chapter, microsatellite loci can be powerful molecular tools for the detection of population genetic structure.

a) In the present Chapter, the same suite of 15 microsatellite loci were employed to test the hypotheses that six natural and three seeded plains zebra (*Equus quagga*) populations from across the sub-region were genetically diverse.

b) These markers as well as population-representative sequences of the rapidly evolving mitochondrial DNA control region were used to test the hypothesis that genetic differentiation within the sub-region is consistent with present subspecific designations.

c) A final goal was to combine the information obtained above into an efficient management plan for the species in the sub-region.

4.2 METHODS AND MATERIALS

Details of the DNA and analytical methods used here are described in Chapter 2.

4.3 RESULTS

4.3.1 DNA isolation and amplification

DNA was isolated from either whole blood or dry salted skins. Some loci amplified more optimally than others, although PCR was successful for all 184 samples.

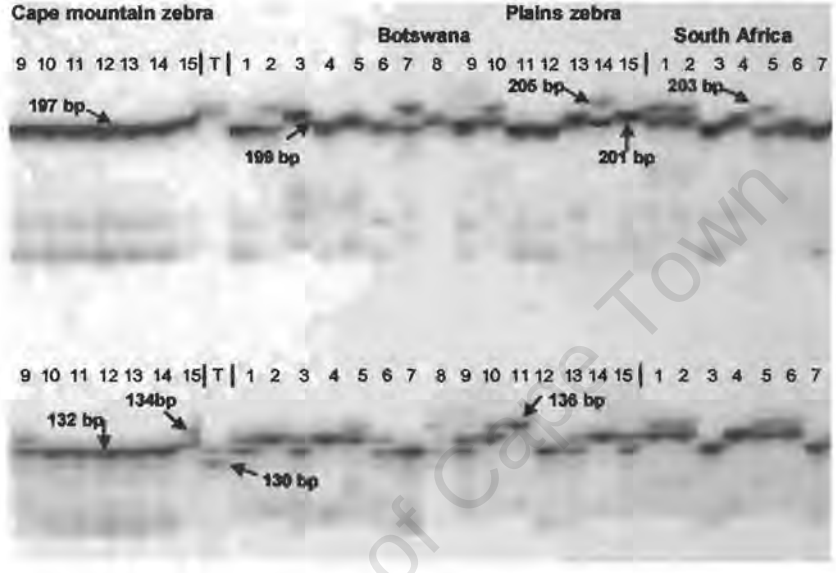

Figure 4.3. Visual comparison of allelic diversity in plains and mountain zebras at two allelically poor microsatellite loci. Locus LEX 52 (Coogle and Bailey, 1997) is pictured above and locus VHL 21 (van Haeringen *et al.* 1998) below. Allele sizes are given in base pairs. Plains zebra samples begin at T, which is the single plains zebra sample from an unknown location in Tanzania.

High genetic variation in plains zebra populations relative to Cape mountain zebras is immediately apparent at loci with both low (Fig. 4.3) and high (Fig. 4.4) allelic diversity.

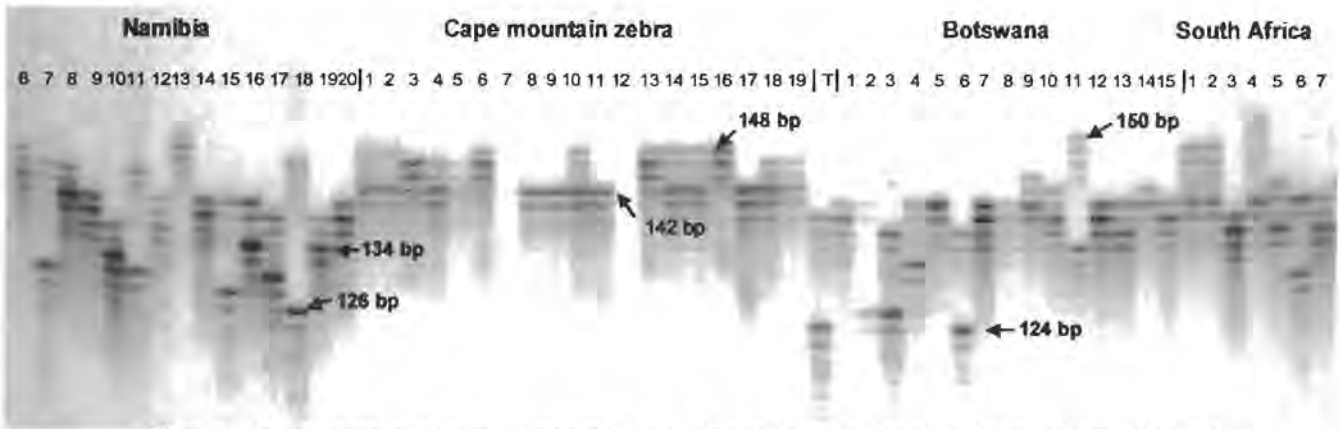

Figure 4.4. Variation between populations of southern African plains zebra (labelled Namibia, Botswana and South Africa) and various Cape mountain zebras at HTG 7 (Ellegren *et al.* 1992), an allelically diverse microsatellite locus. T, is the single plains zebra sample from an unknown location in Tanzania.

4.3.2 Inter-locus disequilibrium

After 1000 Markov chain simulations, the null hypothesis of statistically significant association between loci was rejected at the 99% level ($p > 0.01$) for all pair-wise comparisons. All loci were therefore assumed to be unlinked for this data set.

4.3.3 Nuclear microsatellite genetic variation and population structuring

4.3.3.1 Genetic variation

Observed heterozygosity (H_O) and allelic diversity (A, the mean number of alleles per locus) were significantly higher in plains zebra populations than in either subspecies of mountain zebra (Plains zebra *vs* HMZ: H_O, $p < 0.001$; A, $p < 0.01$; *vs* CMZ: H_O, $p < 0.001$; A, $p < 0.01$; Student's two sample t-test). The mean number of alleles per locus present in the entire plains zebra metapopulation (10.23) was significantly higher than the equivalent value in mountain zebras (6.23, $p < 0.01$) and in horses (6.07, $p < 0.01$) when tested with a Student's two sample t-test. The mean number of alleles was not significantly different between mountain zebra and the horse ($p > 0.05$). Horse allelic data were obtained from the literature for a minimum of 20 unrelated individuals for HMB1 (Binns *et al.* 1995) to over 100 horses of different breeds for UCDEQ 505

(Eggleston-Stott *et al.* 1997). This suggests that horse isolated microsatellites may be more variable or at least as variable as in other equid species.

Consistently high variation was observed in the free ranging plains zebra populations of Zimbabwe, Botswana and Namibia as well as in the Lowveld group of South Africa (Table 4.2). Large differences between H_O and H_E exist in all but the Botswana population. This is a surprising result as all free ranging populations are large and assumed to be in mutation-drift equilibrium. In all the above cases, $H_O < H_E$, indicating a multilocus heterozygote deficit.

Table 4.2. Nuclear genetic variation in plains zebra (*Equus quagga*) based on 15 microsatellite loci. N_F denotes the founder number for seeded populations; N_S, the number of animals sampled in each population; N_A, the number of alleles; N_P, the number of polymorphic loci in each population. The mean number of alleles per locus (A) is corrected for sample size using the jack-knifing procedure. H_O, observed proportion of heterozygotes in each population; H_E, unbiased proportion of expected heterozygotes based on allele frequencies calculated according to Nei (1978). metapop., metapopulation

Population	N_F	N_S	N_A	N_P	A	Heterozygosity (H)	
						H_O	H_E
Western Zimbabwe	-	39	131	15	5.58	0.6624	0.7661
Zambezi Valley	-	8	84	15	5.37	0.6930	0.7621
Namibia	-	23	115	15	5.50	0.6716	0.7868
Botswana	-	15	110	15	5.64	0.7578	0.7663
Lowveld	-	33	122	15	5.70	0.7066	0.7950
KwaZulu metapop.	-	66	74	15	3.88	0.6288	0.6277
Hluhluwe-Umfolozi	-	23	71	15	3.89	0.6716	0.6606
Vernon Crookes	19	19	61	15	3.45	0.6007	0.5886
Albert Falls	12	17	51	14	3.06	0.5702	0.5194
Harold Johnson	8	7	49	15	3.27	0.7019	0.6174
TOTAL		**184**					

Mean diversity was significantly lower in the KwaZulu populations, (H_O, $p < 0.05$; A, $p < 0.001$; Student's two sample t-test) and comparable in magnitude to free ranging Hartmann's mountain zebra. The lower diversity is consistent with that expected in a bottlenecked population and accurately reflects the mass slaughter of game during the

nagana campaigns of the first half of the 20th century. The seeded KwaZulu populations generally contained less diversity than the founding Hluhluwe-Umfolozi stock with the exception of Harold Johnson, the smallest population with a higher H_O. As this population was founded only 20 years ago and the population size has not changed since founding, it would appear that $H_O > H_E$ is a direct result of founding by animals from different family groups.

4.3.3.2 Hardy-Weinberg equilibrium

Only three of a possible 135 population-locus exact tests were significant ($p < 0.01$) after sequential Bonferroni correction (Rice, 1989). Before this, 10 population-locus exact tests showed a significant departure ($p < 0.01$) from Hardy-Weinberg equilibrium and 16 tests were significant at $0.01 < p < 0.05$ (Table 4.3). Fifteen of the 26 failed tests were associated with the free-ranging Northern populations. These reflect a heterozygote deficit, where $H_O < H_E$ for all Northern populations (Table 4.2). However, highly significant global tests showed that each locus and population was in overall HWE.

Table 4.3. Exact tests for departure from Hardy-Weinberg equilibrium (HWE) in plains zebra populations. P values indicate the probability of rejecting the null hypothesis of random association of alleles, that is, HWE. The null hypothesis cannot be rejected at $p > 0.05$ (**) or $p > 0.01$ (*). The null hypothesis of HWE can be rejected for $p < 0.01$. Population/locus combinations still significantly deviating from HWE after sequential Bonferroni correction at $p < 0.01$ were denoted $p < 0.01$**B**. In one instance, no test was possible (-) due to monomorphism across a population. pop., population

Population\ Locus	HTG3	HTG5	HMB1	HTG7	HTG9	HTG 11	HTG 14	HTG 15	LEX 52	VHL 47	LEX 20	UCDEQ 505	AHT 21	TKY 273	COR 014	Global pop.
Western Zimbabwe	*	**	p<0.01	**	**	**	**	p<0.01**B**	**	**	**	*	**	**	p<0.01**B**	**
Zambezi Valley	**	**	*	**	**	**	**	**	**	**	**	**	*	**	**	**
Namibia	**	**	*	**	**	**	**	**	**	**	p<0.01	*	**	**	p<0.01	**
Botswana	**	*	**	**	**	**	*	**	*	**	**	*	**	**	**	**
South Africa	p<0.01**B**	**	*	**	**	*	**	p<0.01	**	*	p<0.01	**	**	**	p<0.01	**
Umfolozi-Hluhluwe	**	**	p<0.01	**	**	**	**	**	**	**	**	**	**	**	**	**
Vernon Crookes	**	**	**	**	**	**	**	**	*	**	**	**	**	**	**	**
Albert Falls	**	**	**	-	**	**	**	**	**	**	**	**	**	**	**	**
Harold Johnson	*	*	**	**	**	**	**	**	**	**	**	**	**	**	**	**
Global locus	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**

4.3.3.3 Analysis of molecular variance (AMOVA)

Genetic structure in plains zebras was determined by the significance of the distribution of the variance components of the data. After 10 000 permutations, the entire plains zebra metapopulation shows moderate but significant genetic structuring, accounting for 9.99% of the variance in the data set (Table 4.4). The two-group scenario separating the KwaZulu populations from all other populations received highest support. The higher Φ_{ST} value for this grouping is a consequence of the loss of genetic variation from the KwaZulu metapopulation and the subsequent differentiating effect of genetic drift. The two subspecific groups scenario (*Chapmani* and *Antiquorum*) received the lowest multi-group support. The separation on the basis of Northern (Zimbabwe, Botswana and Namibia) and Southern (South Africa) groups was also not strongly supported, accounting for only 5.92% of the variation in the data. None of the three group scenarios received high support.

Table 4.4. Analysis of nuclear molecular variance in plains zebra populations. The fixation index Φ_{ST} (equivalent to θ, Weir and Cockerham, 1984) and the percentage of the variation that is distributed among groups are given for different grouping scenarios. ***, significant with $p < 0.001$; **, significant with $p < 0.01$; *, significant with $p < 0.05$; **NS**, not significant. Nam, Namibia; Bots, Botswana

	Combined metapopulation	
	Φ_{ST}	%Va
1 Group	0.09989***	9.99***
2 Groups: *Chapmani* and *Antiquorum*	0.10584***	1.18^{NS}
: Northern and Southern	0.12247***	5.92*
: All populations and KwaZulu	0.14653***	11.11**
3 Groups: Zimbabwe and Nam-Bots and S.Africa	0.11091***	4.06^{NS}
: Zimbabwe and Nam-Bots-Lowveld and KwaZulu	0.11687***	7.93**

4.3.3.4 Exact tests of population differentiation

Pair-wise tests for population differentiation showed that all pairs of populations contained significant levels of allelic heterogeneity (Table 4.5). The two Zimbabwean populations and Hluhluwe-Umfolozi and Harold Johnson were significant only at $0.001 < p < 0.01$, indicating a greater degree of similarity between these pair-wise comparisons. The magnitude of pair-wise differentiation is investigated in the next Section.

Table 4.5. Exact pair-wise tests of differentiation between plains zebra populations. P-values are given for each pair-wise comparison. ***, significant with $p < 0.001$; **, significant with $p < 0.01$; * significant with $p < 0.05$, **NS**, not significant ($p > 0.05$)

	Western Zimbabwe	Zambezi Valley	Namibia	Botswana	Lowveld	Hluhluwe -Umfolozi	Vernon Crookes	Albert Falls
Zambezi Valley	**							
Namibia	***	***						
Botswana	***	***	***					
Lowveld	***	***	***	***				
Hluhluwe-Umfolozi	***	***	***	***	***			
Vernon Crookes	***	***	***	***	***	***		
Albert Falls	***	***	***	***	***	***	***	
Harold Johnson	***	***	***	***	***	**	***	***

4.3.3.5 Pairwise population structuring

Both estimates of population structuring return a range of values from very low and non-significant (θ, 0.0087; *Rho*, 0.0044) to high and significant (θ, 0.2259; *Rho*, 0.3316). The two population pairs that returned the least significant exact tests of differentiation show correspondingly low estimates of population structuring.

Moderate to low values were returned for pair-wise comparisons among populations in the north of the sub-region (upper left box, Table 4.5), indicating high relatedness. Among the Northern group, Zambezi Valley and Western Zimbabwe populations are most similar. However, the lack of marked genetic differentiation among the five non-KwaZulu populations implies that the close association between Western Zimbabwe and the Zambezi Valley is more likely to be a result of close spatial proximity. Observed values show that the Damara zebra population of Botswana is closely associated with the Chapman's zebra population of Western Zimbabwe but not with the Zambezi Valley. The Namibian population appears more closely related to the Botswana population than to any other Northern population. This suggests a pattern of isolation by distance (IBD) across the semi continuous belt of protected wilderness between Etosha National Park in Namibia and Mana Pools National Park in Zimbabwe. Dispersal over the Zambezi Escarpment, therefore, is not inhibited in plains zebras as it is for the artiodactyl giraffe (*Giraffa camelopardalis*) and for wildebeest (*Connochaetes taurinus*), neither of which occur in the Zambezi Valley (Mann, 1990).

The moderate to low structuring between the Hluhluwe-Umfolozi-derived KwaZulu populations (bottom right box, Fig. 4.5) reflects their recent demographic history.

However, a high level of structuring was found between KwaZulu and Northern populations, corroborating the earlier findings of AMOVA (Table 4.4). The lower genetic diversity of KwaZulu populations (Table 4.2) implies that the observed differentiation may have occurred as a result of recent population bottlenecking and accelerated drift. However, differentiation between KwaZulu and the Lowveld sample was intermediate, being high for *Rho* but low for θ estimates.

Table 4.6. Pair-wise fixation indices between plains zebra populations. Values for the F_{ST} estimator θ are given above the diagonal and values for the R_{ST} estimator *Rho* are given below the diagonal. **NS**, not significant $p > 0.05$; *, significant at $p < 0.05$; **, significant at $p < 0.01$; ***, significant at $p < 0.001$, after 10 000 permutations. The box in the upper left quadrant encloses the Northern populations and the box in the lower right quadrant encloses the KwaZulu-derived populations.

	Western Zimbabwe	Zambezi Valley	Namibia	Botswana	Lowveld	Hluhluwe -Umfolozi	Vernon Crookes	Albert Falls	Harold Johnson
Western Zimbabwe		0.0087 **NS**	0.0352 ***	0.0189 **	0.0287 ***	0.1104 ***	0.1361 ***	0.1775 ***	0.1135 ***
Zambezi Valley	0.0087 **NS**		0.0407 **	0.0465 **	0.0274 *	0.1346 ***	0.1666 ***	0.2259 ***	0.1534 ***
Namibia	0.0592 ***	0.0679 *		0.0279 ***	0.0330 ***	0.0989 ***	0.1391 ***	0.1708 ***	0.1071 ***
Botswana	0.0379 *	0.0732 *	0.0438 **		0.0285 ***	0.1126 ***	0.1586 ***	0.1933 ***	0.1319 ***
Lowveld	0.0160 **NS**	0.0243 **NS**	0.0358 **	0.0484 **		0.0665 ***	0.0908 ***	0.1275 ***	0.0681 ***
Hluhluwe-Umfolozi	0.1832 ***	0.2274 ***	0.1067 ***	0.2030 ***	0.1184 ***		0.0610 ***	0.0620 ***	0.0291 **
Vernon Crookes	0.1921 ***	0.2500 ***	0.1423 ***	0.2235 ***	0.1187 ***	0.0383 *		0.0815 ***	0.0609 ***
Albert Falls	0.2857 ***	0.3316 ***	0.2218 ***	0.3162 ***	0.1964 ***	0.0801 ***	0.0441 *		0.0881 ***
Harold Johnson	0.1279 ***	0.1818 **	0.094 **	0.1772 ***	0.0796 *	0.0044 **NS**	0.0432 **NS**	0.1005 **	

4.3.3.6 Genetic distance

The three distance statistics that were used elucidated varying degrees of population genetic structure in southern African plains zebra. The Northern populations and the consolidated Lowveld samples were completely resolved by D_S and D_C with bootstrap support > 50% (Figs 4.5A and 4.5B). In all three cases Western Zimbabwe and the Zambezi Valley were grouped together with not less than 75% bootstrap support. The distance separating these two populations is generally no larger than the distance separating the Northern populations from the Lowveld group. This corroborates previous results from AMOVA (Table 4.4) and fixation indices (Table 4.6) that failed to

detect subspecific structuring between Damara and Chapman's zebra phenotypes. D_S and $(\delta\mu)^2$ placed Botswana as the next most related population, but D_C grouped Botswana with Namibia with 61% bootstrap support. The Lowveld group was found to be intermediate between the Northern group and the KwaZulu populations in all cases. Bootstrap resampling support among KwaZulu populations was low except in the case of $(\delta\mu)^2$. Harold Johnson was most closely related to the Hluhluwe-Umfolozi source.

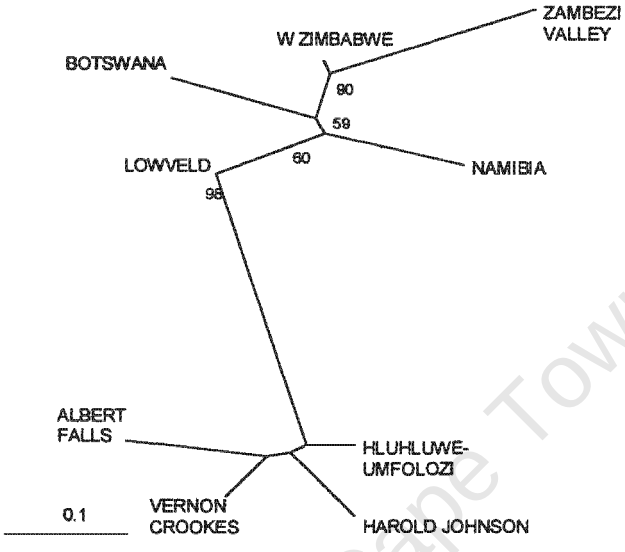

Figure 4.5A. D_S, standard genetic distance, Nei (1972)

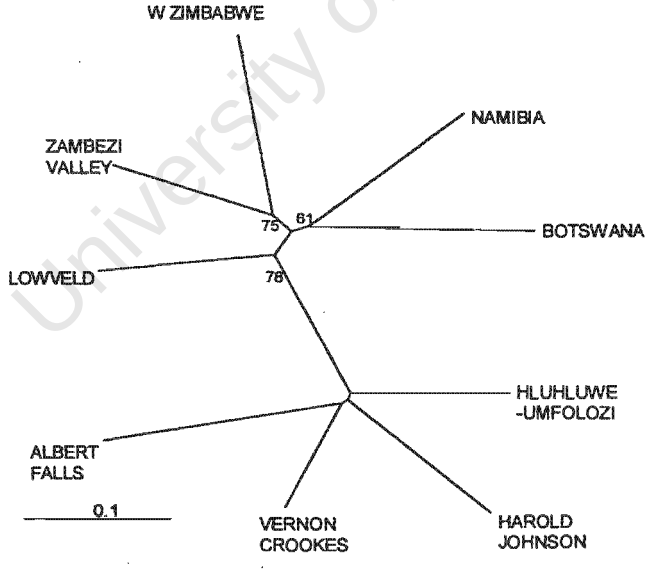

Figure 4.5B. D_C, Cavalli-Sforza and Edwards (1967)

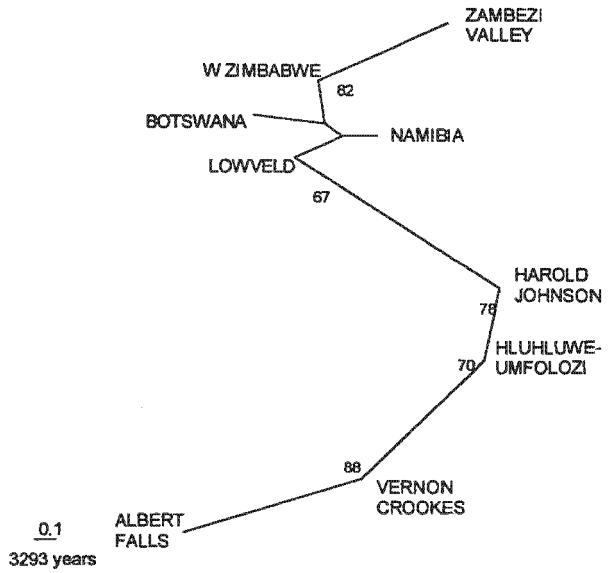

Figure 4.5C. $(\delta\mu)^2$, Goldstein *et al.* (1995)

Figure 4.5. Unrooted neighbour joining trees of three genetic distance measures between populations of plains zebras. Nodal significance was determined by 1000 bootstrap replicates. W, Western

4.3.3.7 Isolation by distance

As all KwaZulu populations are descended from the Hluhluwe-Umfolozi source, the inclusion of them as separate entities is not likely to yield any further useful information. These populations were therefore conflated.

The D_S and D_C distance statistics detected significant isolation by distance ($p < 0.05$) but $(\delta\mu)^2$ did not. The higher variance of the $(\delta\mu)^2$ statistic is probably responsible for this lack of support. The D_C estimate achieved the highest correlation co-efficient (Fig. 4.6B and Table 4.7). The observed distribution of genetic heterogeneity in the plains zebra data thus conforms to that which would be expected as a result of isolation by distance. This therefore appears to be the evolutionary mechanism that effects population structure in both southern African zebra species.

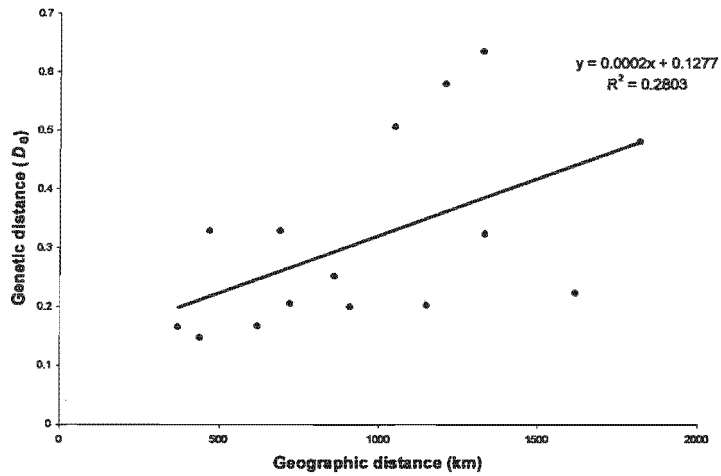

Figure 4.6A. D_S, standard genetic distance, Nei (1972)

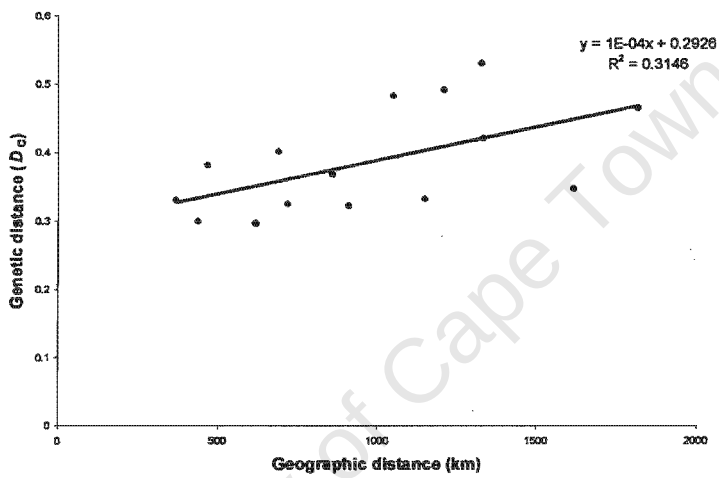

Figure 4.6B. D_C, Cavalli-Sforza and Edwards (1967)

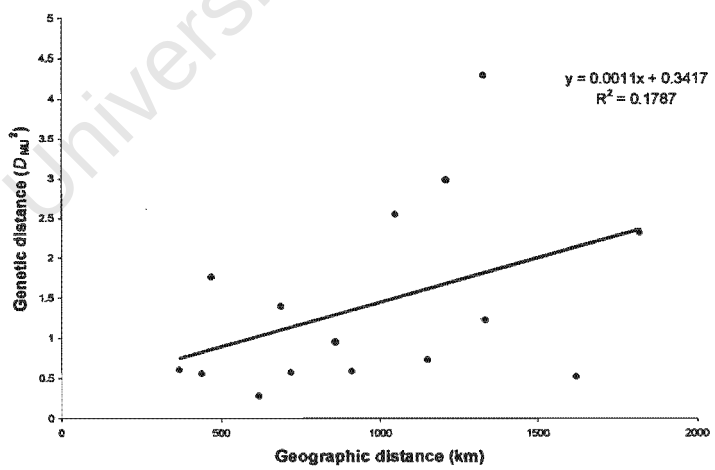

Figure 4.6C. $(\delta\mu)^2$, Goldstein *et al.* (1995)

Figure 4.6. Correlation of geographic distance with three measures of genetic distance in plains zebra populations.

Table 4.7. Test for isolation by distance among plains zebra populations using three distance statistics (D_S, D_C and $(\delta\mu)^2$).

	Coefficient of regression (R^2)	P (any correlation > observed correlation)	Slope of regression line	Y-intercept
D_S	0.2803	0.04880	0.00019444	0.1276517
D_C	0.3146	0.04940	0.00009737	0.2925613
$(\delta\mu)^2$	0.1787	0.18210	0.00111371	0.3416910

Within population genetic diversity (H_E) was negatively but not significantly correlated with all measures of genetic distance (Table 4.8). As in the mountain zebra data, the most severely affected statistic was D_S but unlike the previous data, $(\delta\mu)^2$ was also strongly correlated despite its high variance. The lack of significance is most likely to be associated with the high genetic diversity in most plains zebra populations. The analysis also predicts the level of genetic diversity required in each population for the existence of panmixia across the sub-region (X intercept, Table 4.8).

Table 4.8. Correlation of within-population genetic diversity with three genetic distance statistics (D_S, D_C and $(\delta\mu)^2$).

	Coefficient of regression (R^2)	Significance	Slope of regression line	Y-intercept	X-intercept H_E at distance=0
D_S	0.8389	0.10300	-4.38343567	3.6144307	0.8246
D_C	0.7712	0.08040	-1.92432908	1.8385102	0.9555
$(\delta\mu)^2$	0.8215	0.06490	-30.43494754	24.3360552	0.7996

4.3.3.8 Divergence times

Assuming that 2.05×10^{-4} mutations per generation is correct, the most distant of plains zebra populations in the southern African sub-region diverged no more than an estimated maximum of 108 417 – 234 139 years ago (95% confidence limits of highest value, Table 4.9). This is very similar to the 95 096 – 242 792 years estimated from the same microsatellite loci as the maximum divergence time range for mountain zebra populations. However, as this study only looks at a regional subset of the entire plains zebra metapopulation, the estimated coalescence time for this data set may be of little value, particularly if historical radiation into the sub-region occurred from different sources north of the Zambezi River.

The five highest divergence times result from pair-wise comparisons involving the KwaZulu metapopulation, reinforcing the genetic distinctiveness of this gene pool.

Times since splitting of Northern populations are all less than a maximum of 36 179 – 76 343 years, underlining the relative genetic similarity of these populations.

Table 4.9. Divergence times in years with 95% confidence limits, assuming a generation time of 16.5 years, dating the time of split between plains zebra populations, based on the $(\delta\mu)^2$ distance statistic.

	Western Zimbabwe	Zambezi Valley	Namibia	Botswana	Lowveld
Zambezi Valley	24428 ±9417				
Namibia	30263 ±14367	50546 ±15534			
Botswana	22939 ±9457	56261 ±20082	22255 ±6680		
Lowveld	13602 ±5473	40083 ±12274	22537 ±9980	24348 ±5795	
KwaZulu	102220 ±32598	171278 ±62861	89543 ±24871	114977 ±32598	65074 ±20565

4.3.3.9 Interspecific divergence

In order for microsatellite genetic distance to be useful at the interspecific level, the possibility of a constraint on maximum and minimum allele sizes must be investigated. Although only results graphed in black may be directly comparable among the three equid species, allelic distributions were similar (Fig. 4.7), with a greater range of allele classes in plains zebra. Taking all of the zebra data into account (grey and black bars, Fig. 4.7), the mean number of repeats in plains zebra was not significantly higher than that in mountain zebra (Table 4.10). This was further corroborated by a modal value of 11 repeats for both zebra species. No mean values were significantly different when all three species were tested against each other, although both the mean and modal repeat size was highest in horses (Table 4.10). Despite a lower genetic variability in horses (see Section 4.3.3.1), which is probably associated with horse domestication and consequent inbreeding, the data support the previous work of Bowcock *et al.* (1994) and Forbes *et al.* (1995) which showed that allele size tends to be larger in the species from which the microsatellite loci were isolated.

Also, the fewer copies of large and small repeats indicate a mechanism of microsatellite evolution by which alleles with a small number of repeats tend to get larger and alleles with a large number of repeats tend to get smaller. In horses, the allelic distribution seems unimodal and the mean repeat size is approximated by the

modal value. This may have resulted from the fact that the allelic data presented for horses were derived independently for each locus, from different numbers of horses (see Section 2.5.2.4). The horse data may therefore be regarded as being largely qualitative.

The repeat size distributions (Fig. 4.7) show clearly that repeat size is constrained between four and 24 repeats in mountain zebras, three and 25 repeats in plains zebras and seven and 23 repeats in horses. The proportion of homoplasic allele classes is therefore expected to be high as alleles of the same size may not be identical by descent. The calculated range of 113 335 – 184 969 years for the time since *E. zebra* and *E. quagga* diverged is less than the maximum intraspecific values obtained for both species (242 792 and 234 139 years for mountain zebra and plains zebra respectively). The observed interspecific divergence time estimate is therefore an underestimate caused by the loss of linearity in the $(\delta\mu)^2$ statistic.

Table 4.10. Mean and modal microsatellite repeat size in three equid species. Means were tested by a Student two sample t-test, assuming a normal distribution and unequal variances. Comparisons with the horse were restricted to 14 loci as these were the only data available. **NSH**, not significant ($p > 0.05$) when compared to the mean in horses; **NSM**, not significant ($p > 0.05$) when compared to the mean in mountain zebra; **NSP**, not significant ($p > 0.05$) when compared to the mean in plains zebra; n/a, not applicable

	Mountain zebra	Plains zebra	horse
Mean repeat size (averaged across 15 loci)	12.7 ^{NSP}	13.2 ^{NSM}	n/a
Mean repeat size (averaged across 14 loci)	12.0 $^{NSP, NSH}$	12.7 $^{NSM, NSH}$	14.4 $^{NSM, NSP}$
Modal repeat size	11	11	13

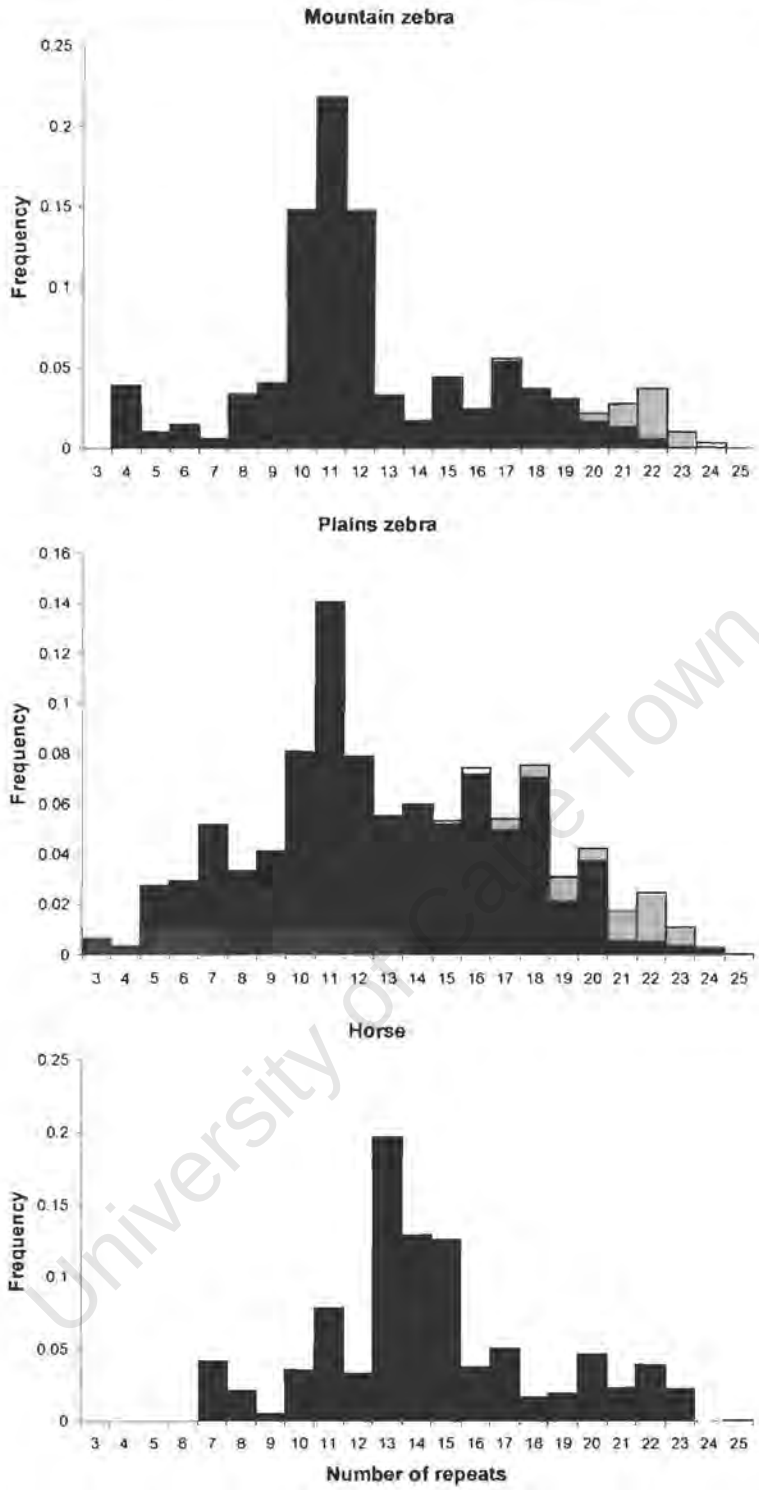

Figure 4.7. Allelic distributions showing the similarity in repeat length that results from constraints in allele size in mountain zebra, plains zebra and the horse. Bars in black describe the distribution of alleles in the 14 microsatellite loci for which horse allelic data were available. Grey bars show the added data for the remaining locus (AHT 21) in zebras. The horse distribution is standardised relative to the total zebra data.

4.3.3.10 Gene flow

A dissimilar number of migrants among all populations per generation was returned for the two N_em calculations (4.714326, N_em_{PA}; 1.90029, N_em_{RST}). Both N_em estimates however, detected homogenising gene flow between the closely related, free ranging Northern populations of approximately >1 (Table 4.11). The lowest of these values (0.9762, N_em_{PA}) corresponds to the Namibia and the Zambezi Valley, two populations on opposite ends of the sampling range that spans the sub-region from east to west.

The correlation between the two estimates of gene flow was weak, but significant (R^2 = 0.2724, $p < 0.05$, Mantel test, 10 000 permutations). Three of the four highest values were N_em_{RST} estimates associated with the Lowveld population. As this is strictly not a population but a combination of samples of Kruger National Park origin, the N_em values obtained are most likely to describe historical gene flow through the South African Lowveld.

The loss of diversity and drift that has affected the KwaZulu populations was again manifested in low gene flow estimates in associated pair-wise comparisons. That the lowest of these values were between KwaZulu and Northern populations and the highest level of gene flow into KwaZulu appears to be from the South African Lowveld, suggests that the observed genetic structuring may be, at least partly historical, and not human mediated.

Table 4.11. Effective number of migrants (N_em) between the plains zebra populations in southern Africa. Values for the N_em estimate obtained from *Rho* (N_em_{RST}) according to (Ciofi and Bruford, 1999) are given below the diagonal, and values obtained by the private alleles method (N_em_{PA}) (Slatkin, 1985) are given above the diagonal.

	Western Zimbabwe	Zambezi Valley	Namibia	Botswana	Lowveld	KwaZulu
W Zimbabwe		2.0642	1.5184	2.9203	1.9040	0.8794
Zambezi Valley	14.3112		0.9762	1.5334	1.467	0.2011
Namibia	1.9851	1.7164		2.069	1.164	0.6865
Botswana	3.1742	1.5825	2.7259		1.769	0.6989
Lowveld	7.7011	5.0155	3.3669	2.4570		1.1091
KwaZulu	0.5169	0.3844	0.8095	0.4322	0.9023	

4.3.3.11 Population assignment tests

Despite a lack of heterogeneity within Northern and KwaZulu groups, multilocus likelihood assignment tests provided strong support for genetic structuring within these groups (Table 4.12). This high population fidelity led to the estimation of strong local structuring into Northern, Southern and KwaZulu groups. KwaZulu was however, the only local group with 100% correct assignment. Intra-population assignment was weak for Western Zimbabwe and the Lowveld, suggesting high historical gene flow through these populations. Despite its high sensitivity, the multilocus population test also failed to assign more than 95% of Chapman's zebras, thus corroborating evidence from AMOVA (Table 4.4) and fixation indices (Table 4.6).

Table 4.12. Percentage of multilocus plains zebra genotypes correctly assigned to their populations or localities of origin. Locality and subspecific values are written in bold.

Population	% of individuals correctly assigned		
	Population	**Locality**	**Subspecies**
Northern		**99**	
Western Zimbabwe	90		
Zambezi Valley	100		
Namibia	100		
Botswana	100		
Southern		**99**	
Lowveld	91		
Kwazulu		**100**	
Hluhluwe-Umfolozi	100		
Vernon Crookes	95		
Albert Falls	100		
Harold Johnson	100		
E. q. chapmani			**94**
E. q. antiquorum			**100**

4.3.3.12 Principal component (PC) analysis

After 1000 permutations, only PC one was significant, accounting for 54.4% of the inertia in the microsatellite data ($p < 0.01$, abscissa axis, Fig. 4.8). PC two described 10.9% of the inertia ($p > 0.1$, ordinate axis, Fig. 4.8). The total inertia in the plains zebra data set described by the two-dimensional Figure 4.8 is therefore 65.3%, but only separation along the abscissa axis will be considered here as significant. Local genetic structuring in southern African plains zebra is not supported as the Lowveld population clusters closer to the Northern group along the abscissa axis. The *E. q. chapmani* grouping occurs only along the ordinate axis and is not significant. PCA provides further support for the grouping of Hluhluwe-Umfolozi with its satellite populations and the genetic separation of these KwaZulu populations from others in the sub-region.

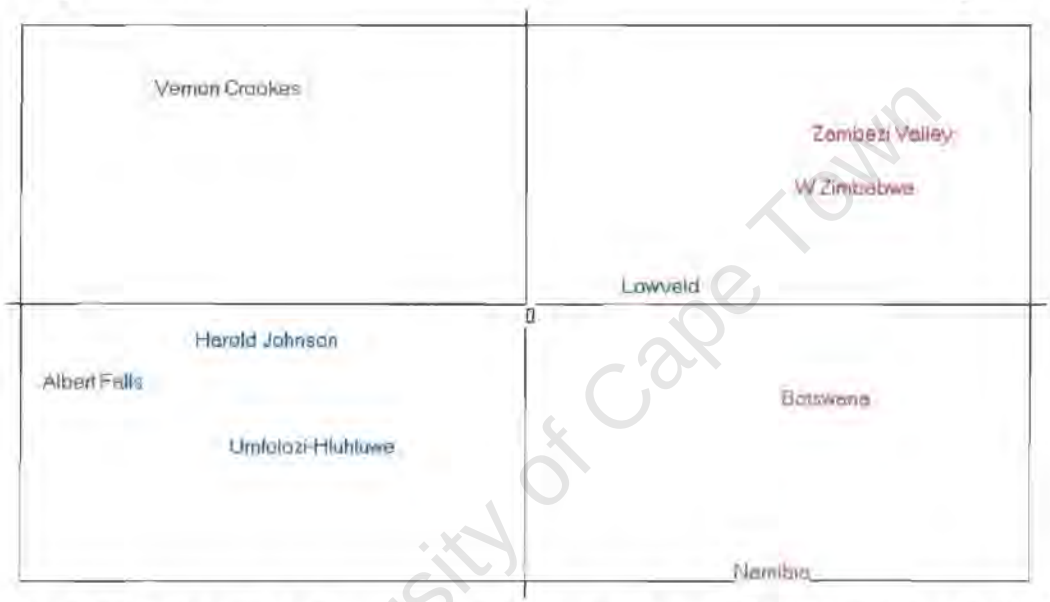

Figure 4.8. Plains zebra populations subjected to principal component analysis. Populations are plotted against the first two principal components which describe the majority of the variance in the data. Northern populations are in red and pink; Southern populations in blue. *Equus quagga chapmani* are in red; *E. q. antiquorum* in pink and blue.

4.3.3.11 Effective population size

Genetic estimates of N_e were high (>1000, Table 4.13) for all *E. quagga* populations, reflecting the high levels of genetic diversity observed previously (Table 4.2). Census estimates for all Northern populations and direct estimates of N_e were not higher than genetic estimates in Namibian and Zimbabwean populations, suggesting that extant plains zebras in these areas are not able to maintain their present levels of genetic

diversity. In the case of the Lowveld group, most populations were seeded from Kruger National Park stock. The Kruger National Park population is estimated at 25 250 and is probably large enough to maintain the observed genetic diversity. Since each Lowveld population is small however, the current level of diversity is expected to decrease without intensive management. A similar but more pronounced trend is apparent in the smaller KwaZulu populations where even the Hluhluwe-Umfolozi population of 3000 individuals is not large enough to prevent the loss of diversity due to drift. The values calculated assuming the IAM are consistently lower than values assuming the SMM.

Table 4.13. Comparison of extant and effective population size N_e in plains zebras. a, Hack *et al.* 2002; b, Department of Environmental Affairs, Namibia, c, Department of Wildlife and National Parks, Botswana; d, G. Castley, pers. comm.; e, Bowland *et al.* 2001; f, determined by the formula $1/N_e=1/4$ (1/Br M +1/Br F), assuming 81.8% of any population consists of breeding adults, with a sex ratio of 1:1.38 (Smuts, 1974). KrNP, Kruger National Park; SMM, step-wise mutation model; IAM, infinite alleles model

Population	Extant Population Size	Effective Population size (N_e)		
		Direct Estimate [f]	Assuming SMM	Assuming IAM
Western Zimbabwe	5710 [a]	4551	9606	3772
Zambezi Valley	3165 [a]	2522	6732	3012
Namibia	5000 [b]	3985	11415	4196
Botswana	29 955 [c]	23 873	8451	3481
Lowveld (KrNP)	25 250 [d]	20 123	12 303	4392
Hluhluwe-Umfolozi	3000 [a]	2391	4251	2224
Vernon Crookes	110 [e]	88	2726	1633
Albert Falls	50 [e]	40	1861	1235
Harold Johnson	9 [e]	7	2731	1635

4.3.3.13 Population bottlenecking

Garza and Williamson's (2001) *M* ratio was consistent with population history data for all but the Zambezi Valley population which returned an *M* estimate of less than 0.81 (Table 4.14). This is consistent with the discrepancy between direct and genetic estimates of N_e (Table 4.13) which also alludes to a reduction in size for the Zambezi Valley population. All other Northern populations as well as the Lowveld group showed no signs of a recent reduction in size and even minimum values are above the 0.81 cut off.

By contrast, all KwaZulu populations as well as the combined KwaZulu metapopulation appear to have undergone recent population bottlenecking. This is consistent with known historical records and is also confirmed by the reduced genetic diversity observed in these populations. Interestingly, the source population, Hluhluwe-Umfolozi, does not return the highest *M* value. The lowest estimated *M* corresponds to Harold Johnson, the smallest population with the smallest number of founders, despite relatively high observed genetic diversity (see Table 4.2).

The heterozygote excess test failed to detect population bottlenecking in all but the smallest KwaZulu population ($p < 0.05$) under the assumption of the SMM. When the IAM was assumed, all populations were significantly bottlenecked ($p < 0.05$) despite their demographic histories. These inconsistencies render this method unworkable in the reliable detection of recent reductions in population size.

Table 4.14. Comparison of two tests for population bottlenecking in all plains zebra populations. *M* < 0.81 (Garza and Williamson, 2001) or p-values for heterozygote excess < 0.05 (Luikart *et al.* 1998) indicates population bottlenecking. Minimum *M* calculated as *M* minus variance of *M*. metapop., metapopulation

Population	Garza & Williamson's *M* (minimum *M*)	Variance of *M*	Number of non-monomorphic loci N_P	P-value: Heterozygote excess SMM	IAM
Western Zimbabwe	0.875 (0.864)	0.011	15	0.98721	0.00003
Zambezi Valley	0.787 (0.746)	0.041	15	0.78940	0.01508
Namibia	0.850 (0.825)	0.025	15	0.51102	0.00003
Botswana	0.852 (0.838)	0.014	15	0.83487	0.00513
South Africa	0.854 (0.824)	0.030	15	0.38077	0.00002
Hluhluwe-Umfolozi	0.752 (0.696)	0.056	15	0.21060	0.00050
Vernon Crookes	0.780 (0.717)	0.063	15	0.70026	0.00066
Albert Falls	0.734 (0.677)	0.057	14	0.29150	0.01477
Harold Johnson	0.714 (0.660)	0.054	15	0.01508	0.00021
KwaZulu metapop.	0.761 (0.706)	0.055	15	0.31934	0.00011

4.3.4 Mitochondrial haplotype structuring

4.3.4.1 Sequence variation

The 22 plains zebra control region lineages consisted of 20 haplotypes. Of the 756 bp sequenced, 12.8% of the sites (97 sites) were polymorphic when the *E. caballus* outgroup was included in the data set and 8.1% (61 sites) were variable when the outgroup was excluded (Table 4.15). Sixty two site patterns were found among the 20 *E. quagga* haplotypes with a high transition/transversion ratio of 8.55. Two haplotypes were shared between four individuals. Both shared haplotypes were South African, one of which was exclusive to the Lowveld population and the other found in both the Lowveld and KwaZulu populations. The high proportion of unique haplotypes was reflected in high haplotype diversity (0.991 ± 0.017). All haplotypes were found to be closely related despite high estimated substitution rate heterogeneity for the data set, α = 0.09. Incorporating the α parameter, nucleotide diversity was calculated according to the Kimura 2-parameter model (Kimura, 1980) π^1= 0.02271 ± 0.00862, and by ML, π^2= 0.02168 ± 0.01175. In both cases, nucleotide diversity was higher among plains zebra haplotypes than among mountain zebra haplotypes. Within the southern African sub-region, nucleotide diversity was marginally lower at π^1= 0.02166 ± 0.00824 and π^2= 0.01969 ±0.01075.

Table 4.15. Polymorphic sites in 756 bp of plains zebra of control region sequence.

```
                10        20        30        40        50        60
            ----|----|----|----|----|----|----|----|----|----|----|----|-
BOTSWANA1   AAGCTTTACTTACACTGAAGTGAGATTCTAAAAGCCCAACATTCAGAAACTGTTAAAAGGCAAGAA
BOTSWANA2   -AG--CCATTTATACTGGGATGTGATTCTAAGAGTTCAATACTTAGCAGTCGTCAAAAGGCAAGGA
BOTSWANA3   AAG--TCACTTACACTGAAGTGAGATTCTAAGAGCCCAATACTCAGAAGTTGCTGAAAGATAGAGG
BOTSWANA4   AAG--TCACTTACACTGAAGTGAGATTCTAAGAGCTCAATACTCAGAAGTTGCTGAAAGGTAAAAG
NAMIBIA1    A-G--TCACTTACACTAAAGTGAGATTCTAAGAGCCCAAC-CTCAGAAGTTGTTGAAAGGTAAGGG
NAMIBIA2    AAG--TTACTTACACTGAAACGAGATTCTAAAAGCCCAACATTCAGAAGCTGTTAAAAGGCAAGAA
NAMIBIA3    AAG--TCACTTACACTGAAGTGAGATTCTGAGAGCCCAGCACTCAGAAGTTGTTAAAAGGCAAAAG
KWAZULU1    AAG--TCACTTACACTGGAGTGAGATTCTAAGAGCTCAATACTCAGAAGTTGCTGAAAGGTAGGGG
KWAZULU2    AAG--TCACTTACACTGGAGTGAGATTCTAAGAGCTCAATACTCAGAAGTTGCTGAAAGGTAGAGG
KWAZULU3    AAG--TCACCTACACTGAAATGAGATTCTAAGAGCCCAATACTCAGAGGTTGTTGAGAGGTAAGAG
LOWVELD1    -AG--TCATTCACACTGAAATATAGTTTTGAGAGCTTTACACCCGGCAATTGCTGAAAGGTAAAAG
LOWVELD2    AAG--TCACTTACATTGAAATAAGATTCTAAGAGCCCAATACTCAGAAGTTCTCAGAAGGCAAGAG
LOWVELD3    AAG--TCACTTACACTGAAGTGAGATTCTAAGAGCCCAATACTCAGAAGTTGCTGAAAGGTAAAAG
LOWVELD4    AAG--TCACTTACACTGGAGTGAGATTCTAAGAGCTCAATACTCAGAAGTTGCTGAAAGGTAGAGG
LOWVELD5    AAA--TCACTTACGCTGAAGTGAGATTCTAGGGACCCAATACTCAGAAGCTGTTGAAAGGTGAGAG
LOWVELD6    AAA--TCACTTACGCTGAAGTGAGATTCTAGGGACCCAATACTCAGAAGCTGTTGAAAGGTGAGAG
TANZANIA    -AG--CC-TTTGTACTGAAATGTGACCTTAAGAGTTCAATACTTAGCAGTTGTCAAAAGGTAAGAG
WZIMBABWE1  AAG--TCACTTACACTGAAGTGATATTCTAAGGGCCCAATACTCAGAAGTTGCTGAAAGGTAAAAG
WZIMBABWE2  AAG--TCACTTACACCGAAATGAGATTTTAAAAGCCCAACACTCAGAAGCTGTTAAAGAACAAAGG
ZAMBEZI1    AAG--TCACTTACACTGAAGTGAGATTCTAGGAGCTCAATACTCAAAAGTTGTTGAAAGGCAAAGG
ZAMBEZI2    -AG--CC-TTTGTACTGGGATGTGATTTCAAGAGTTCAATACTTAGCGGTTGTCAAAAGGTAAGGG
ZAMBEZI3    AAG--TCACTTACACTGAAGTGAGATTCTAAGAGCTCAATACTCAGAAGTTGCTGAAAGGCAGGAG
```

4.3.4.2 Mismatch distribution

The multimodal observed distribution of pair-wise differences between *E. quagga* control region haplotypes (Fig. 4.9) is not consistent with that predicted by Rogers and Harpending's (1992) sudden expansion theory (expected curve, Fig. 4.9). Instead, multimodality is usually associated with the stochastic shape of gene trees of populations in demographic equilibrium (Schneider *et al.* 2000).

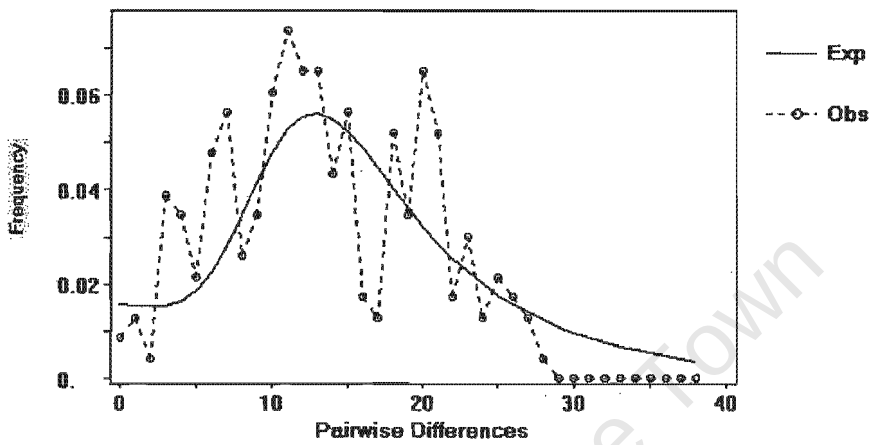

Figure 4.9. Mismatch distribution of pair-wise differences between plains zebra mitochondrial haplotypes. Observed (Obs) pair-wise differences differ to that expected (Exp) if populations underwent sudden expansion.

4.3.4.3 Analysis of molecular variance (AMOVA)

AMOVA was used to test nine different grouping hypotheses (Table 4.16) across 22 plains zebra sequences that spanned seven populations. The only significant result after 10 000 permutations was the percentage of variation between Northern, Southern and Tanzanian groups ($0.01 < p < 0.05$). As no Tanzanian population was typed with nuclear microsatellites, a direct comparison between nuclear and mitochondrial AMOVA results was not possible. The highest Φ_{ST} value was returned when the variance in all populations was compared to that of the Tanzanian sample. Neither this value nor any other Φ_{ST} was significant. Molecular variance of the control region provided no support for local or subspecific genetic partitioning of plains zebra mitochondrial lineages.

Table 4.16. Analysis of haplotypic molecular variance in plains zebra populations. The fixation index Φ_{ST} and the percentage of the variation (%Va) that is distributed among groups are given for different grouping scenarios. *, significant with $p < 0.05$; **NS**, not significant. Nam, Namibia; Bots, Botswana

Grouping scenarios	Combined metapopulation	
	Φ_{ST}	%Va
1 Group : All populations	0.01884 ^{NS}	1.88 ^{NS}
2 Groups: Northern (incl. Tanzania) and Southern	0.03010 ^{NS}	2.77 ^{NS}
: All populations and Tanzania	0.31123 ^{NS}	33.30 ^{NS}
: All populations and KwaZulu	-0.01772 ^{NS}	-5.21 ^{NS}
3 Groups: *Chapmani* and *Antiquorum* and *Boehmi*	0.05066 ^{NS}	6.55 ^{NS}
: Northern and Southern and Tanzania	0.05633 ^{NS}	10.76*
: All populations and KwaZulu and Tanzania	0.08219 ^{NS}	10.33 ^{NS}
4 groups : Zimbabwe and Nam-Bots and S.Africa and Tanzania	0.03186 ^{NS}	7.37 ^{NS}
: Zimbabwe and Nam-Bots-Lowveld and KwaZulu and Tanzania	0.02762 ^{NS}	2.98 ^{NS}

4.3.4.4 Isolation by distance

Despite a stronger correlation coefficient between geographic distance and between-population genetic distance (R^2 = 0.5649, Fig. 4.10) than in mountain zebras, the correlation was not significant ($p = 0.1068$, Mantel test, after 10 000 permutations). The high proportion of unique haplotypes and high within-population genetic variation may explain this lack of significance. The inclusion of only one *E. q. boehmi* sample without precise geographic information (sample attributed to the Serengeti National Park) may also confound the observed result. Although more samples are required from more locations across the *E. quagga* distributional range in order to perform an adequate test for IBD, a positive trend between geographic and genetic distance is found, though it is not significant.

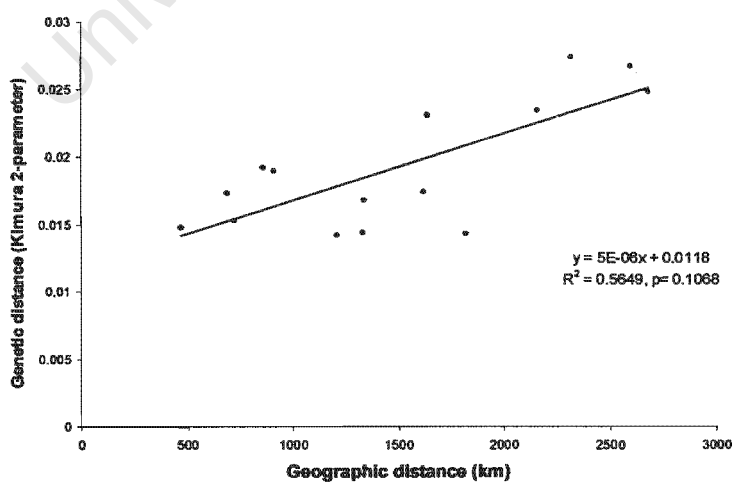

Figure 4.10. Isolation by distance among plains zebra mitochondrial haplotypes.

4.3.4.5 Phylogenetics

The quartet puzzling algorithm (Strimmer and von Haeselar, 1996), assuming the HKY model of nucleotide substitution and the intrageneric rate heterogeneity estimate of α = 0.12 (Oakenfull *et al.* 2000), produced an unresolved maximum likelihood (ML) tree (Fig. 4.11). Despite a significant difference in likelihood scores that favoured a non-clock-like tree ($p < 0.01$, χ^2 test), the tree assuming a molecular clock was used, thus enabling the computation of coalescence time.

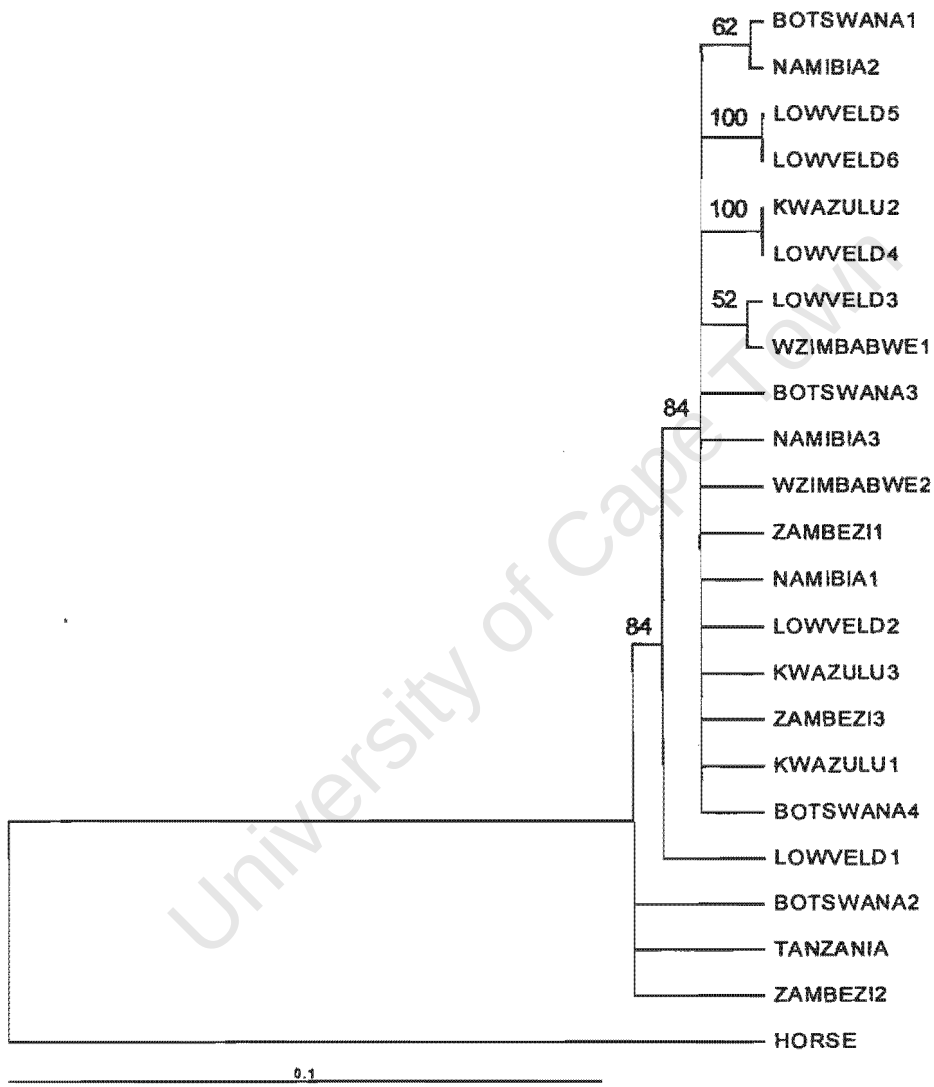

Figure 4.11. 50% majority rule consensus maximum likelihood phylogenetic hypothesis for 22 plains zebra control region haplotypes. The horse (*E. ferus*) was chosen as the outgroup. Maximum likelihood branch lengths were calculated assuming a molecular clock. Quartet puzzling values are given for internal branches with greater than 50% support.

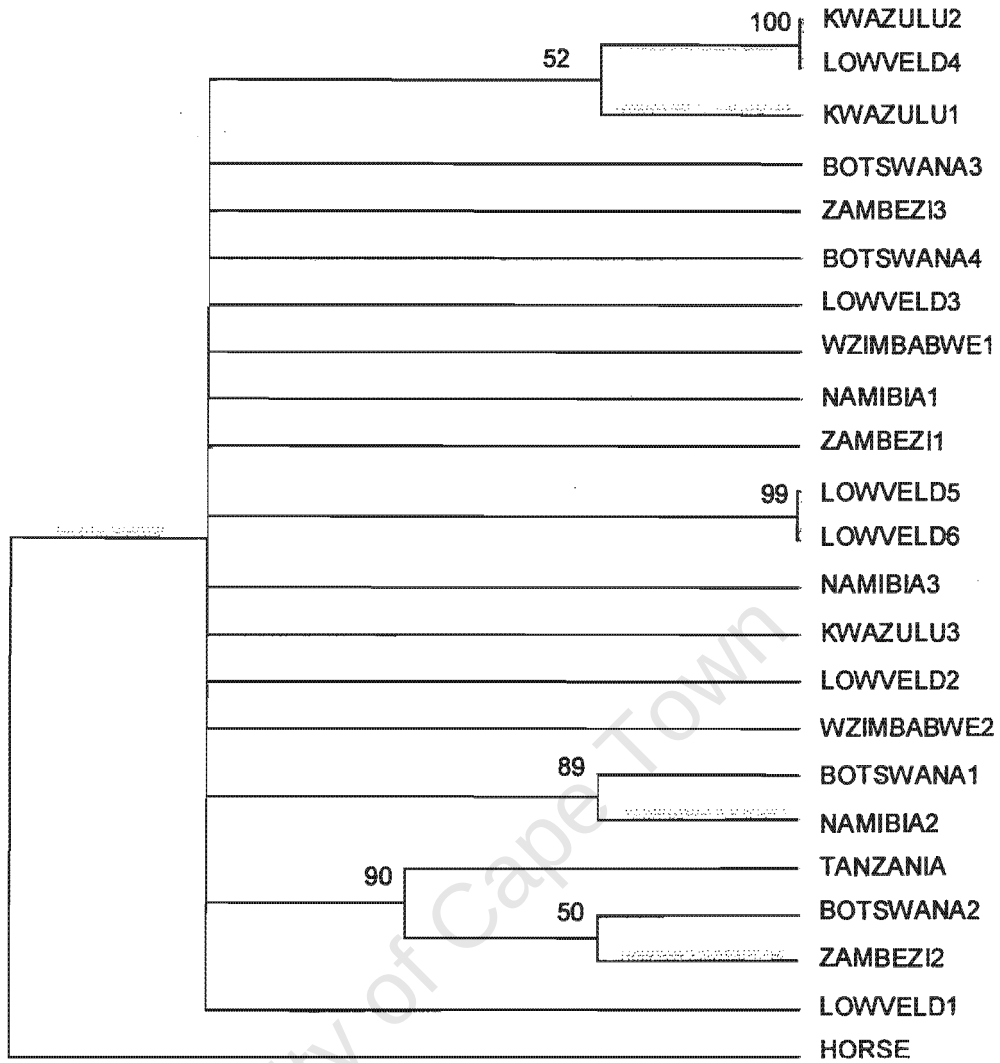

Figure 4.12. 50% majority rule consensus maximum parsimony phylogenetic hypothesis for 22 plains zebra control region haplotypes. The horse (*E. caballus*) was chosen as the outgroup. Only nodes supported by more than 50% of 500 bootstrap replicates are shown.

Subjecting the aligned sequence data to maximum parsimony (MP) analysis resulted in a topologically similar but less resolved tree (Fig. 4.12). The lack of resolution is most likely to be caused by the high observed haplotypic diversity and low nucleotide diversity, the latter being represented by only 33 parsimony informative characters. A control region sequence from the horse (*E. caballus*) was used to root both trees. No phylogeographic structure was apparent in either tree hypothesis. Small maximum likelihood branch lengths relative to the outgroup sequence (Fig. 4.11), however, imply a relatively recent coalescence of *E. quagga* control region lineages.

Subspecific haplotypic structuring was also unsupported. Neither *E. q. chapmani* (Western Zimbabwe and Zambezi Valley), *E. q. antiquorum* (Namibia, Botswana and South Africa) nor *E. q. boehmi* (Tanzania) haplotypes were monophyletic. Phylogenetic analyses partitioned the Tanzania sequence separately with one representative of each of the two other subspecies, indicating higher genetic divergence between the East African sample and the majority of the southern African haplotypes. This supports the survey of Oakenfull *et al.* (2000) where incomplete division between East African (*E. q. boehmi*) and southern African (*E. q. antiquorum* and *E. q. chapmani*) plains zebra haplotypes was observed. This is also consistent with the between-population IBD analysis (Fig. 4.10) where *E. q. boehmi* was related most distantly to all southern African populations.

4.3.4.6 Coalescence

Both estimates of divergence between the plains zebra data set and the *E. caballus* outgroup were corrected for substitution rate heterogeneity using $\alpha = 0.12$, calculated by Oakenfull *et al.* (2000) for control region haplotypes across the genus. The average nucleotide divergence, D_{xy} (Nei, 1987), was calculated using Kimura's (1980) 2-parameter model as 0.17444 (± 0.06435). Genetic divergence calculated by maximum likelihood analysis, assuming the HKY model of nucleotide substitution, was 0.12740 (± 0.04955). The first reliable fossil evidence for *E. quagga* was dated at 0.7 Mya (Eisenmann, 1992) which Oakenfull *et al.* (2000) used to calibrate their control region/12S rRNA molecular clock of the Equidae, estimating the existence of a common *E. quagga-E. caballus* ancestor at a maximum of 2.3 Mya. Using $\mu = D_{xy}/2T$, nucleotide substitution rates of 3.79×10^{-8} substitutions/site/year (3.79%/Myr), or (by ML analysis) 2.77×10^{-8} substitutions/site/year (2.77%/Myr) were calculated for the data. Substituting for the new μ values and using both overall estimates of nucleotide diversity within plains zebras, 299 604 years (95% CI: 185 883 – 413 324 years) by the Kimura 2-parameter model (Kimura, 1980) or 391 336 years (95% CI: 179 188 – 603 484 years) by ML, were calculated as the time necessary for the coalescence of studied *E. quagga* mitochondrial lineages. Coalescence estimates for southern African haplotypes were marginally lower at 285 752 years (Kimura 2-parameter estimate, 95% CI: 177 044 – 394 459 years) and 355 415 years (ML estimate, 95% CI: 161 372 – 549 458 years). This indicates that either the nucleotide diversity in southern African plains zebras is close to that for the entire species or that the one Tanzanian sample was more similar to the southern African haplotypes by chance.

4.4 DISCUSSION

4.4.1 Sampling Effects

The two smallest population sample sizes were for Harold Johnson Nature Reserve (n=7) and the Zambezi Valley (n=8). In the case of the former, the number of sampled animals was only two short of the total population size. Given that only eight animals originally founded the population, the sample size obtained for this study is probably adequate to describe the genetic complement of the Harold Johnson population. The number of plains zebra in Zimbabwe's Zambezi Valley is estimated at 3165 animals (Hack *et al.* 2002). It is therefore unlikely that this small sample size provides an accurate estimate of population allele frequencies. Despite this, this sample was not significantly differentiated from its neighbouring population of Western Zimbabwe and only returned a deviant result for the Garza and Williamson (2001) test for population bottlenecking.

4.4.2 Statistics

4.4.2.1 Genetic distance

Owing to a lack of genetic differentiation, relatedness among plains zebra populations in the north of the southern African sub-region could not be consistently elucidated. Concordant with all three distance measures was the placement of the Lowveld population as being most closely related to the KwaZulu group and the large relative distance separating KwaZulu populations from all others. Although within population genetic diversity delivered a non-significant correlation co-efficient when plotted against genetic distance, distance statistics did not perform as well as with the mountain zebra data set. This is perhaps related to the relative dearth of genetic diversity in mountain zebra populations. The highly diverse plains zebra populations may contain too many homoplasic back mutations which adversely affect genetic distance. Poor performance may also have resulted from the sampling of only a subset of the species range and thus not obtaining a complete overview of the distribution of genetic heterogeneity across the species.

4.4.2.2 Population assignments and their usefulness in determining structure

The insensitivity of this method as a means of determining population structure was again apparent in the plains zebra data set as it supported hypotheses that were rejected by other analyses presented earlier. This demonstrates conclusively the unsuitability of this indirect assessment of population structure.

4.4.2.3 PCA

Multidimensional partitioning of plains zebra populations underlines again the utility of the PCA approach not only to identify the areas of heterogeneity in microsatellite data but also to impart statistical robustness to an otherwise qualitative graphical plot. The validity of these results was corroborated by AMOVA (Table 4.4), fixation indices (Table 4.6) and genetic distances (Figure 4.5).

4.4.2.4 Bottlenecking

Based on results from both the plains and mountain zebra data sets, it may be concluded that the Garza and Williamson (2001) test for population bottlenecking is most reliable when sample size is larger than ten and when all loci used in the test are polymorphic. It was also found that the use of at least 15 loci returned results that were consistent with population history records.

Conversely, the more complex approach that relies on the detection of heterozygote excess consistently returned results that were in contrast to known demographic records. This may be linked to the social phenomenon of behavioural isolation (Blake *et al.* 1981) that may lead to the detection of a heterozygote deficit in a random sample of naturally occurring equids. The failure of this method with the extensive data sets of both southern African zebra species may be a direct consequence of this and suggests that heterozygote excess may not be a feature of equid population bottlenecks.

4.4.2.5 Coalescence estimates

The outside limit for the coalescence of plains zebra haplotypes was 603 484 years, almost 100 000 years longer than the mountain zebra equivalent, implying a greater diversity in plains zebra control region lineages. This may be overtly illustrated in the high phenotypic diversity of plains zebra across its range and especially in southern Africa. The mean coalescence times of 299 604 and 391 336 fall within the estimate of

700 000 years when *E. quagga* first appeared on the fossil record. The inclusion of one Tanzanian sample in the present analysis increased mean coalescence time estimates by 13 852 and 35 921 years respectively. As paraphyly of southern and East African groups has previously been observed (Oakenfull *et al.* 2000), this increase may not be sufficient to describe adequately the coalescence of the entire species. The ML estimate of nucleotide divergence for the data set of Oakenfull *et al.* (2000) which included seven East African (*E. q. boehmi*) and southern African (*E. q. chapmani and E. q. antiquorum*) haplotypes, was 0.0298, using $\alpha = 0.12$. Substituting this value using the *E. quagga-E. caballus* calibrated substitution rate, delivers a more accurate mean coalescence time of 537 906 years for the entire species.

When accelerated drift was invoked to explain the high divergence estimates associated with the KwaZulu population, the highest nuclear estimate of population divergence time (234 139 years) fell within the lower limits of all the mitochondrial estimates. However, the highest mean nuclear coalescence estimate of 171 278 years fell short of the lowest mitochondrial estimate of 177 044 years, indicating a departure from the SMM and the loss of linearity of $(\delta\mu)^2$. This has been explained by constraints in allele size (see Section 4.3.3.9) for interspecific divergence. Paetkau *et al.* (1997) found that even the $(\delta\mu)^2$ statistic may plateau as early as 3000 generations since divergence. With a generation time of 16.5 years, this would correspond to 49 500 years. This study shows that constraints in allele size affects genetic distance on a short time scale (>100 000 years), so rendering microsatellite data inappropriate, even for intraspecific coalescence estimates.

4.4.2.6 AMOVA

Neither the nuclear nor the mitochondrial level AMOVA was successful in elucidating phylogeographic structure in plains zebra populations. The only discrepancy between the two analyses was the irregular distribution of variance associated with the differentiation of the KwaZulu gene pool from other populations that was detected at microsatellite loci but not in mitochondrial DNA sequences. That KwaZulu is differentiated only by hypervariable, recombining markers indicates that the differentiating phenomenon occurred relatively recently.

4.4.3 Heterozygote deficit

The multilocus heterozygote deficit observed for most Northern populations (Table 4.2) probably results from the biasing effect of a few loci as global tests could not demonstrate significant multilocus Hardy-Weinberg disequilibrium (Table 4.3). Loci appeared to exhibit disequilibria at random as no locus exhibited significant disequilibria for all Northern populations.

The phenomenon of allozyme heterozygote deficiency was observed by Bowland *et al.* (2001) for the KwaZulu data set used in this study and by Blake *et al.* (1981) in feral donkeys. This deficiency is thought to result from behavioural isolation as a consequence of polygynous equid social organisation. As herd stallions sometimes have lifelong breeding rights on female harems, breeding and non-breeding (bachelor) herds may become socially segregated. Rare alleles could be protected from extinction by localised fixation in these socially isolated groups, thus leading to the Wahlund effect and a consequent heterozygote deficit (Bowland *et al.* 2001)

4.4.4 Population structuring and gene flow

As with mountain zebras, there was support for the IBD hypothesis of population structuring in southern African plains zebras at the nuclear microsatellite level. In all cases the correlation coefficient was weak due to the lack of marked differentiation between Lowveld and Northern populations and the distribution of a significant proportion of the genetic heterogeneity in the plains zebra microsatellite data between the KwaZulu group and all others.

Gene flow was calculated to be close to or higher than one migrant per generation among all non-KwaZulu populations suggesting genetic homogeneity across that part of the sub-region. Non-differentiation among these southern African populations is not surprising given that genetic diversity is high, that there is no evidence for population bottlenecking and that there are few putative barriers to inhibit gene flow. Among these barriers may be the Zambezi River, the Zimbabwean Plateau, from which plains zebras seem to be excluded, the Kalahari, where lack of water may preclude zebra habitation and the Drakensberg Mountains over which gene flow is likely to be infrequent owing to the rough terrain.

Gene flow across the protected area just south of the Zambezi-Chobe-Kavango system was found to occur widely as the populations from this area all show moderate to low pair-wise population differentiation. The area north of this system is unlikely to promote gene flow of comparable magnitude due to the patchy distribution of protected areas and the high incidence of poaching in Zambia and Angola. Occasional south to north trans-Zambezi movement may occur on either side of Lake Kariba as the river itself does not act as a dispersal barrier to large ungulates, especially in the dry season. The Zimbabwe Plateau may also be dismissed as a potential barrier to gene flow. Zambezi Valley and Lowveld, two populations on either side of the plateau, showed low levels of pair-wise structuring (θ, $p < 0.05$; *Rho*, $p > 0.05$) and comparably high gene flow estimates ($N_e m_{RST} = 5.0155$, $N_e m_{PA} = 1.467$). The presence of a large population (n=9000, Hack *et al.* 2002) in the aeolian Kalahari sands of central Botswana's Makgadikgadi Game Reserve indicate that low annual precipitation does not limit *E. quagga* distribution. The Central Kalahari would thus not have been a significant hindrance to gene flow. High differentiation was found in pair-wise comparisons of KwaZulu and all but the Lowveld population. This differentiation may be ascribed to increased drift as a result of the drastic reduction of the KwaZulu population in the first half of the 20th century. These reduction events are also consistent with *M* values obtained for the Hluhluwe-Umfolozi stock population. But if increased drift was fully responsible for the observed high differentiation between KwaZulu and Northern populations, then why are Lowveld plains zebras not as highly differentiated from the KwaZulu population?

KwaZulu plains zebras are noted for their higher than average incidence of reduced striping and lightening of the hindquarters (Rau, 1978). This race, previously called the Wahlberg's zebra (*E. q. wahlbergi*) by Antonius (1951), may thus have existed historically in relative isolation. This may also have contributed to the detection of marked population structuring with the far-off Northern populations of Namibia, Botswana and Zimbabwe and intermediate structuring with the neighbouring Lowveld population as gene flow across the Drakensberg may have been attenuated. The bulk of the gene flow into KwaZulu must have been through the Lowveld. The observed population structuring between KwaZulu and other populations in the sub-region may therefore be a combination of human mediated and natural drift. The lack of resolution in control region sequences makes this hypothesis difficult to substantiate, although the only shared haplotype in the mitochondrial data set originated from Lowveld and KwaZulu populations. KwaZulu plains zebras may thus have remained in relative

geographical isolation from populations west of the Drakensberg while being supplemented with haplotypes from the Lowveld.

4.4.5 Phylogeography

The results here show clearly the shortcomings of an intra-regional phylogeographic analysis. As the present study did not have the means to sample plains zebras from outside southern Africa, a significant phylogeographic assessment was obtained at only the nuclear level. However, weak IBD was detected among a limited number of control region haplotypes with high within-population haplotype variation. If the net genetic distance between populations (that is, only the component of the overall distance that is between populations) is tested against geographic distance, the correlation is strong and highly significant ($R^2 = 0.7871$, $p = 0.0068$, Mantel test, after 10 000 permutations). This suggests that phylogeographic structure in plains zebra may mirror that found in mountain zebras despite a non-linear sub-Saharan distribution.

Oakenfull *et al.* (2000) also found that inadequate sampling left questions unanswered. The division between southern and East African haplotypes observed by them was not detected here as only one East African sample was available. It is therefore possible that an evolutionary history similar to that of the mountain zebra exists for the plain zebra where widely separated populations appear more dissimilar to each other. If this is the case, East and southern African gene pools may be regarded as separate entities with common haplotypes occurring at low frequency owing to the clinal distribution of gene frequencies that results from isolation by distance. This trend has been demonstrated qualitatively ($R^2 = 0.5649$) here by the available haplotypic data. However, to test this hypothesis rigorously, samples distributed intermediately in Zambia, the southern DRC and northern Mozambique, as well as more East African samples, must be obtained and analysed at both the nuclear and mitochondrial levels.

Gene frequencies were not significantly different between the well-sampled neighbouring southern African subspecies (Damara and Chapman's zebras) at either the nuclear or the mitochondrial levels. The complete lack of subspecific structuring in mitochondrial data suggests the existence, perhaps until relatively recently, of large, free ranging populations in close association with each other across most of the sub-

region. In the absence of an allopatric barrier to gene flow, the present subspecific delineation between Chapman's and Damara zebras must be questioned.

4.4.6 Management strategy

4.4.6.1 Northern populations

Veterinary cordon fences in Botswana seem to do little to prevent west-east gene flow to and from Namibia. Subsistence poaching is rife in both parks and intervening areas. Despite this, the high summer rainfall and the availability of permanent water still support large plains zebra populations. Genetic diversity in these populations is high and animals still free-range along the semi-continuous network of protected areas south of the Kavango, Chobe and Zambezi Rivers. For these populations to be regarded as panmictic, pair-wise genetic distance measures would have to approach zero. In distance-heterozygosity plots, the extrapolated H_E at distance zero ranges from 0.7996 for $(\delta\mu)^2$ to 0.9555 for D_C, with D_S being intermediate. The actual expected heterozygosity of the northern samples (Table 4.2) only approaches the lower of these X-intercepts, indicating that these populations are not strictly panmictic. Although the data support the IBD hypothesis, the low level of population sub-division in the northern area may have led to low observed IBD correlation co-efficients. The exchange of genes between populations in this part of southern Africa was high and the plains zebras of the northern populations may therefore be regarded as a single metapopulation. Present national management schemes will only maintain diversity at present levels if these populations are strictly protected and if gene flow is able to occur as freely as in the past. There is already evidence to suggest that population sizes have been reduced to numbers that may not maintain high genetic diversity in the long-term. As this population spans at least three countries, its management also requires regular communication among the appropriate national conservation authorities.

Although present in large numbers, as migratory bulk grazers, zebras are integral members of these sensitive ecosystems and protection may therefore be effected in a variety of ways. Stricter anti-poaching laws and an increased number of anti-poaching staff and patrols needs to be implemented. The maintenance of watering points in dryer areas such as Chobe National Park in northern Botswana and in Hwange National Park in western Zimbabwe must continue. Human-made barriers such as game fences must be kept to a minimum. As in all sensitive systems, there is also the chance of

over-management as seen in the fenced Etosha National Park, Namibia where over 3000 zebra were culled between 1995 and 1998 due to overpopulation (Dept. of Environmental Affairs, Namibia). The management strategy outlined above, if implemented, is likely to impact positively on all wildlife species in the northern part of the southern African sub-region.

4.4.6.2 Southern populations

Although the Lowveld group appears similar to the Northern populations, the KwaZulu metapopulation has been significantly differentiated. The most likely cause of this differentiation is population bottlenecking and drift over the last 100 years although there is evidence to suggest that KwaZulu plains zebra existed in relative isolation prior to this. Nevertheless, the KwaZulu gene pool does not appear to contain unique genetic material at either level of molecular analysis. The management implications of this are largely positive. There does not appear to be any genetic reason against the movement of plains zebras across the sub-region. This increases the number of potential sources from which prospective game farmers may purchase animals and so enable the easier maintenance of genetic diversity.

Almost 30% (16 410) of South Africa's 55 686 plains zebra exist in populations of less than a hundred on small, sometimes privately owned and usually fenced game farms and reserves (Hack *et al.* 2002). Although the Lowveld population appeared diverse, our sample was obtained from no fewer than nine private farms. Singularly, genetic diversity in each small game farm is expected to be low. As these farms were seeded with animals from the Kruger National Park and surrounding Lowveld game reserves, it is proposed that the level of diversity in the Lowveld sample may represent that existing in the Kruger National Park. A similar situation exists on a number of private and provincial reserves in KwaZulu, most of which were seeded with a small number (<20) of Hluhluwe-Umfolozi zebras. The seeded populations studied here have all shown a decrease in genetic diversity. Because they have only been in existence for the last 25 years, however, these populations still retain moderate levels of genetic diversity, as shown by calculated genetic estimates of N_e. This diversity will be eroded by drift if population size remains small. Bowland *et al.* (2001) modelled the loss of genetic variation in small zebra populations and concluded that in order to minimise the loss of heterozygosity to 10% over 100 years, a reserve should be restocked with a harem every five years for a population with n<10 or every 15 years where n>100. With a

growing population of 3000, the Hluhluwe-Umfolozi Park has not reached its maximum carrying capacity. However, if the population continues to increase at its present rate, more animals will become available for the proposed translocation scheme of Bowland *et al.* (2001). Translocating animals from different locations within southern Africa to the same seeded reserves will also help to maintain diversity in small populations. Regulations preventing the purchase of animals from the same supplier more than once and the strict cataloguing of all animal transactions should be implemented in order to aid this process.

4.4.7 Conclusions

Equus quagga, the most successful of all wild equids, maintains a semi-continuous distribution across the vast expanse of sub-Saharan African grassland. This species emerged from the fossil record 0.7 Mya and the mitochondrial lineages studied coalesce in the range of 179 188 – 603 484 years ago. Despite numerous human induced pressures, this species still thrives in large populations throughout its distributional range. Mitochondrial variation has been detected between populations on either side of this range (Oakenfull *et al.* 2000) but this has not been illustrated by nuclear microsatellites nor have populations intermediate in range been sampled. Consequently, the phylogeography of *E. quagga* is yet to be completely elucidated.

In the southern part of its range, *E. quagga* has achieved its highest phenotypic diversity, perhaps as a result of the diverse number of distinct grassland habitats within the southern African sub-region. This high diversity was mirrored at both the nuclear and mitochondrial levels but despite possible adaptive advantages of the differing phenotypes, genetic diversity was not accordingly distributed. Instead, population structure followed that of isolation by distance where the distribution of genetic heterogeneity is clinal, varying with geographic distance. The Drakensberg Mountains may have been the only possible hindrance to gene flow in the sub-region and thus may have contributed to the IBD effect by reducing gene flow into KwaZulu to only via the Lowveld. The southern African sub-region's two extant subspecific forms are not supported by either level of molecular analysis. This is most significantly emphasised by the microsatellite data where the clinal distribution of genes runs approximately along a north-west/south-east plane, with the same subspecies occurring at each end of the distribution. It is therefore suggested that *E. q. antiquorum* and *E. q. chapmani* should be conflated.

As human encroachment and habitat destruction increase, the southern African metapopulation is likely to become more fragmented and drift will differentiate localities one from another. The first part of this process has already occurred in South Africa where the country's largest wild populations of Hluhluwe-Umfolozi Park and Kruger National Park exist in fenced isolation. There is also the threat of losing diversity by protecting isolated populations in small private, municipal and provincial reserves. The absence of genetic structuring in southern African plains zebra populations provides a useful tool for the maintenance of genetic diversity of the metapopulation because inter-population translocations may be carried out freely without the loss of local population genetic integrity.

Chapter 5: Conclusions

This study has sought to understand and define the structure of populations of two southern African zebra species using molecular methods. The most useful molecular tool in this regard was an array of 15 short tandemly repeated microsatellite loci isolated from horses. Until very recent human mediated habitat fragmentation and exploitation, zebra populations were large, free ranging and relatively linearly distributed along either mountain escarpment or savanna grassland. The microsatellite loci employed in this study showed these factors to have led to a nuclear population structure of isolation by distance for both species. However, the hypervariability of microsatellites that normally make them so useful in population genetics rendered them inadequate at higher levels of classification.

5.1 Peculiarities in equid evolution

The variable control region of equid mitochondrial DNA showed moderate phylogeographic structuring on the basis of isolation by distance for both species, suggesting very similar evolutionary histories for *E. zebra* and *E. quagga*. That more phylogenetic resolution was not obtained, especially among *E. quagga* haplotypes, may bring into question the appropriateness of the control region in assessing equid population structure intraspecifically. However, as the control region contains phylogenetic signal at the terminal nodes of the equid phylogenetic tree (Oakenfull *et al.* 2000), it is unlikely that this region of the mitochondrial genome may be evolving too rapidly for the elucidation of intraspecific genetic structure. Oakenfull *et al.* (2000) have argued that rapid radiation and sympatric speciation events may have occurred once ancestral equids moved into areas where different niches were available. This is consistent with the fossil record and has been described by MacFadden and Hulbert (1988) and MacFadden *et al.* (1999). The effect of one or more rapid radiation events would affect the ability of mitochondrial sequences to completely resolve intraspecific phylogenies as the large number of closely related lineages may not have been extant long enough to phylogeographically assort. The present study has shown that extant mountain and plains zebra mitochondrial lineages evolved from striped stenoid

ancestors recently, a maximum of 507 089 and 603 484 years ago, respectively. Purported subspecies of both mountain and plains zebra may thus have diverged too recently to classify them as evolutionarily distinct entities, especially when gene flow between subspecific demes is unhindered. On the other hand, a lack of resolution in phylogenetic trees in both this study and that of Oakenfull *et al.* (2000) may stem from a high number of missing or unsampled haplotypes. Notwithstanding this, Vila *et al.* (2001) sequenced 616 bp of the same control region fragment in modern horses and observed high haplotype diversity and a lack of phylogenetic resolution similar to that found in the present study. They attributed this to a number of independent domestication events that gave rise to the domestic horse. More reliable structuring of multilocus microsatellite genotypes was explained by female biased breeding (Vila *et al.* 2001). The social structure of free ranging zebras precludes invoking this convenient mechanism to explain the very similar microsatellite and mitochondrial results found in the present study.

It is therefore proposed that the mechanisms of evolution in both mitochondrial and microsatellite DNA in the Equidae are very similar and are perhaps even peculiar to the genus. While it may not be possible to differentiate between species with microsatellite data, the present study, as well as that of Vila *et al.* (2001), demonstrate their superiority over mitochondrial DNA in the determination of intraspecific equid population structure.

5.2 Research gaps

5.2.1 Taxa

Future research should be directed at population structuring in Central and East African plains zebra populations to determine if the distribution of genetic heterogeneity follows a similar pattern to that of populations in southern Africa. Grevy's zebra (*E. grevyi*) was shown to be a sister taxon to the plains zebra, with mountain zebras being more distantly related to both (Oakenfull *et al.* 2000). Grevy's zebra is the rarest of the zebras and is listed as endangered by the IUCN (Hilton-Taylor, 2000), with a world estimate of about 5000 animals (Eliot, 1993). Nothing is known of the population's structure or its genetic status. A similar study to the present one, employing the same molecular techniques, will benefit the conservation management of this species. Furthermore, territoriality in males may have different genetic consequences and the

heterozygote deficiency observed here in free ranging plains zebras may be exacerbated in Grevy's zebra.

5.2.2 Microsatellite evolution

The present study has sampled widely enough within the two southern African zebra species to make the observed data statistically comparable (see Section 4.3.3.9). The horse data presented here, however, were not obtained from the same individuals over all the studied loci. Further investigation into the mechanisms of microsatellite evolution that enforce constraints in allele size and maintain allelic modality in equids is required. Such a study would require a large sample of horses and perhaps donkeys in order to obtain unequivocal allelic distributions for interspecific comparisons.

5.2.3 Alternative molecular research

Variation at the non-neutral major histocompatibility complex (MHC) locus may be regarded as a measure of fitness, thus providing a definitive link to inbreeding depression in genetically depauperate populations. The information in the MHC locus may therefore be of conservation interest in the endangered and threatened equid species and subspecies. In the case of the Cape mountain zebra, MHC typing is unlikely to unearth information that is not already known, as the link between the incidence of equine sarcoid and MHC variation has already been established.

Microarray technology may be of potential use in conservation genetics as it may be able to elucidate differential levels of intraspecific gene regulation between outbred and inbred populations and between wild-caught and captive-bred individuals. This is a novel approach to conservation genetics as it looks at transcriptionally-active genes, thereby linking DNA form and function. The outcome of such a study may therefore help revolutionise the conventional way in which population and conservation genetic questions are addressed.

5.3 Zebra conservation

Although the details of management strategies have been discussed in their appropriate chapters, the areas identified as being of immediate conservation relevance are now summarised here. These include the immediate seeding of Cape mountain zebra populations of mixed stock, the removal of cattle from land in the vicinity of mountain zebra reserves and the regulation of the game fences of private landowners in Namibia that are already starting to hinder natural movements. Plains zebra, being more numerous and more genetically diverse, are of less critical concern. However, it is up to the national conservation authorities of Namibia, Botswana, Zimbabwe and South Africa to maintain their presently large populations of *E. quagga* despite the inevitable pressures exerted by an increasing human population. In South Africa, this problem is intensified by the existence of small, isolated and poorly regulated populations on private game reserves. While ecotourism may be the means by which Africa's wildlife pays for itself, conservation authorities must not allow population genetic diversity to be sidelined in favour of short-term economic gain.

After the lion and the elephant, the zebra is one of the major assets to African tourism (Smuts, 1974). The present study has succeeded in obtaining baseline genetic information for the two zebra species extant in southern Africa and has described the dominant evolutionary forces in the generation of zebra population structure. In this way, a contribution to both applied and fundamental knowledge has been made. It is believed that the aims of this project, described in Chapter One, have been fulfilled. It remains to be seen to what extent this information will be integrated into conservation management programmes for southern Africa's striped equids.

References

Anonymous 1970/71. The Andries Venter research centre. *The Provincial administration of the Cape of Good Hope, Department of Nature Conservation annual report* **27**: 33–38.

Ansell, W. F. H. 1960. *Mammals of Northern Rhodesia.* Government Printer, Lusaka.

Ansell, W. F. H. 1974. Order Perrisodactyla. In: *The mammals of Africa: an identification manual* (eds Meester, J., Setzer, H. W.) **14**: 1 – 14. Smithsonian Institution Press, Washington D. C

Antonius, O. 1951. Die Tigerpferde: die Zebras. *Monogm Wildsäugetiere* **11**: 1 – 148.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K. 1987. *Current Protocols in Molecular Biology* **1**. Wiley Interscience

Avise, J. C. 1994. *Molecular Markers, Natural History and Evolution*. Chapman and Hall, New York.

Ballard, J. W. O., Kreitman, M. 1995. Is mitochondrial DNA a strictly neutral marker? *Trends. Ecol. Evol.* **10**: 485 – 488.

Ballou, J. D., Gilpin, M., Foose, T. J. 1995. *Population Management for Survival and Recovery.* Columbia University Press, New York.

Belkhir, K. 2001. Genetix 4.01, Software for population genetics. Laboratoire Génome, Populations et Interactions, University of Montpelier, Montpelier, France.

Bigalke, R. 1952. Early history of the Cape mountain zebra (*Equus zebra zebra*. Linn). *African Wildlife* **6**: 143–153.

Binns, M. M., Holmes, N. G., Holliman, A., Scott, A. M. 1995. The identification of polymorphic microsatellite loci in the horse and their use in thoroughbred parentage testing. *Br. Vet. J.* **151**: 9 –15.

Blake, J. G., Douglas, C. L., Thompson, L. F. 1981. Spatial variation in transferrin allele frequencies among herds of feral donkeys in Death Valley National Monument, California. *J. Mammal.* **62**: 58 – 63.

Boddaert, P. 1785. *Elenchus Animalium* **1**. C. R. Hake, Rotterdam, pp 174.

Boroughs, D. 2000. Rallying around an extinct antelope. *International Wildlife* **32**. 28 – 35

Bowcock, A., Ruiz-Linares, A., Tomfohrde, J., Minch, E., Kidd, J. R., Cavalli-Sforza L. L. 1994. High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* **368**: 455 – 457.

Bowland, A. E., Bishop, K. S., Taylor, P. J., Lamb, J., van der Bank, F. H., van Wyk, E., York, D. 2001. Estimation and management of genetic diversity in small populations of plains zebra (*Equus quagga*) in Kwazulu-Natal, South Africa. *Biochem. Syst. Ecol.* **29**: 563 - 583.

Broders, H. G., Mahoney, S. P., Montevecchi, W. A., Davidson, W. S. 1999. Population genetic structure and effect of founder events on the genetic variability of moose, *Alces alces*, in Canada. *Mol. Ecol.* **8**: 1309 1315.

Brostrom, H. 1995. Equine sarcoids. A clinical and epidemiological study in relation to equine leucocyte antigens (ELA). *Acta Vet. Scand.* **36**: 223 – 236

Brostrom, H., Fahlbrink, E., Dubath, M. L., and Lazary, S. 1988. Association between equine leucocyte antigens (ELA) and equine sarcoid tumors in the population of Swedish halfbreds and some of their families. *Vet. Immunol. Immunopath.* **19**: 215 – 223

Bruford, M. W., Wayne, R. K. 1993. Microsatellites and their application to population genetics. *Curr. Opin. Genet. Devel.* **3**: 939 – 943.

Bryden, H. A. 1889. *Kloof and karroo: Sport, legend and natural history in Cape Colony, with a notice of the game birds, and of the present distribution of antelopes and larger game.*, Longmans, Green and Co., London and New York

Bucher, K., Szalai, G., Marti, E., Griotwenk, M. E., and Lazary, S. 1996. Tumor-Suppressor gene p53 in the horse - identification, cloning, sequencing and a possible role in the pathogenesis of equine sarcoid. *Res. Vet. Sci.* **61**: 114 - 119

Buri, P. 1956. Gene frequency drift in small populations of mutant *Drosophila*. *Evolution* **10**: 367 - 402.

Butler, G. 1999. Striped body, white belly. *Timbila* **4**: 88 – 95

Cabrera, A. 1936. Subspecific and individual variation in the Burchell zebras. *J. Mammal.* **17**: 89 - 112.

Cann, R. L., Brown, W. M., Wilson, A. C. 1984. Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. *Genetics* **106**: 479 – 499.

Castley, G., Lloyd, P.H., Moodley, Y. 2002. Cape mountain zebra, *Equus zebra zebra* taxon data sheet. *IUCN Conservation assessment management plan.* Randburg, South Africa.

Cavalli-Sforza, L. L., Edwards, A. W. F. 1967. Phylogenetic analysis: models and estimation procedures. *Evolution* **32**: 550 – 570.

Charlesworth, D., Charlesworth, B. 1987. Inbreeding depression and its evolutionary consequences. *Ann. Rev. Ecol. Syst.*: **18**: 237 - 268.

Ciofi, C., Bruford, M. W. 1999. Genetic structure and gene flow among Komodo dragon populations inferred by microsatellite loci analysis. *Mol. Ecol.* **8**: S17-S30

Coogle, L., Reid, R., Bailey, E. 1996. Equine dinucleotide repeat loci LEX015 – LEX024. *Anim. Genet.* **27**: 217 – 218.

Coogle, L., Bailey, E. 1997. Equine dinucleotide repeat loci LEX049 – LEX063. *Anim. Genet.* **28**: 378.

Cornuet J.M. and Luikart G., 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**: 2001 - 2014.

Crow, J. F. 1986. *Basic concepts in population, quantitative and evolutionary genetics.* pp 29 – 69. Freeman, New York.

DeSalle, R., Giddings, L. V., Templeton, A. R. 1986. Mitochondrial DNA variability in natural populations of Hawaiian *Drosophila* I. Methods and levels of variability *in D. silvestris* and *D. heteroneura* populations. *Heredity* **56**: 75 - 85.

Deuter, R., Pietsch, S., Hertel, S., Muller, O. 1995. A method for preparation of faecal DNA suitable for PCR. *Nucleic Acids Research*, **23**: 3800 - 3801.

Dorst, J. and Dandelot, P. 1970. *A field guide to the larger mammals of Africa*. Collins, London.

Eck, R. V. and M. O. Dayhoff. 1966. *Atlas of Protein Sequence and Structure*. National Biomedical Research Foundation, Silver Springs, Maryland.

Edmands, S. 1999. Heterosis and outbreeding depression in interpopulation crosses spanning a wide range of divergence. *Evolution* **53**: 1757 - 1768.

Edwards, A., Civitello, A., Hammond, H. A., Caskey, C. T. 1991. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am. J. Hum. Genet.* **49**: 746 – 756.

Eggleston-Stott, M. L., DelValle, A., Bautista, M., Dileanis, S., Wictum, E., Bowling, A. T. 1997. Nine equine dinucleotide repeats at microsatellite loci UCDEQ136, UCDEQ405, UCDEQ412, UCDEQ425, UCDEQ437, UCDEQ467, UCDEQ487, UCDEQ502, UCDEQ505. *Anim. Genet.* **28**: 370 – 371.

Eisenmann, V. 1992. Origins, dispersals and migrations of *Equus* (Mammalia, Perrisodactyla). *Courier Forsch.-Inst. Senckenberg.* **153**: 161 – 170.

Eizirik, E., Bonatto, S. L., Johnson, W. E., Crawshaw, P. G., Vié, J.-C., Brousset, D., O'Brien, S. J., Salzano., F. M. 1998. Phylogeographic patterns and evolution of the mitochondrial DNA control region in two neotropical cats (Mammalia, Felidae). *J. Mol. Evol.* **47**: 613 - 624.

Eizirik, E., Kim, J., Menotti-Raymond, M., Crawshaw, P. G., O'Brien, S. J., Johnson, W. E. 2001. Phylogeography, population history and conservation genetics of jaguars (*Panthera onca*, Mammalia, Felidae). *Mol. Ecol.* **10**: 65 – 79.

Eliot, J.L. 1993. Horse of a different colour is vanishing. *Natl. Geog.* **184**: 142.

Ellegren, H., Johansson, M., Sandberg, K., Andersson, L. 1992. Cloning of highly polymorphic microsatellites in the horse. *Anim. Genet.* **23**: 133 – 142.

Excoffier. L., Smouse, P., Quattro, J. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes. Application to human mitochondrial DNA restriction data. *Genetics* **131**: 479 – 491.

Feldman, M. W., Bergman, A., Pollock, D. D., Goldstein, D. B. 1997. Microsatellite genetic distances with range constraints: analytic description and problems of estimation. *Genetics* **145**: 207 - 216.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evol.* **17**: 368 - 376.

Ferris, S. D., Sage, R. D., Prager, E. M., Titte, U., Wilson, A. C. 1983. Mitochondrial DNA evolution in mice. *Genetics* **105**: 681 - 721.

Fisher, R. A. 1935. The logic of inductive inference. *J. Royal. Stat. Soc.* **98**: 39-54.

Forbes, S. H., Hogg, J. T., Buchanan, F. C., Crawford, A. M., Allendorf, F. W. 1995. Microsatellite evolution in congeneric mammals: domestic and bighorn sheep. *Mol. Biol. Evol.* **12**: 1106 – 1113.

Garza, J. C., Slatkin, M., Friemer, N. B. 1995. Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. *Mol. Biol. Evol.* **12**: 594 – 603.

Garza, J. C., Williamson, E. G. 2001. Detection of reduction in population size using data from microsatellite loci. *Mol. Ecol.* **10**: 305 – 318.

Girman, D. J., Kat, P. W., Mills, M. G., Ginsberg, J. R., Bomer, M., Wilson, V., Fanshaw, J. H., Fitzgibbon, C., Lau, L. M., Wayne, R. K. 1993. Molecular genetic and morphological analyses of the African wild dog (*Lycaon pictus*). *J. Hered.* **84**: 450 – 459.

Goldstein, D. B., Linares, A. R., Feldman, M. W., Cavalli-Sforza, L. L. 1995. An evaluation of genetic distances for use microsatellite loci. *Genetics* **139**: 463 – 471.

Goodman, S. J. 1997. R_{ST}Calc: A collection of computer programs for calculating unbiased estimates of genetic differentiation and determining their significance for microsatellite data. *Mol. Ecol.* **6**: 881 – 885.

Goodrich, L., Gerber, H., Marti, E., and Antczak, D. F. 1998. Equine sarcoids. *Vet. Clin. N. Amer. Equine Pract.* **14**: 607.

Goudet, J. 1999. PCAgen, version 1.2. Software for Principal Component Analysis (PCA) on genetic data. Institute of Ecology, University of Lausanne, Switzerland.

Gray, J. E. 1824. A revision of the family Equidae. *Zool. J. London*, **1**: 241 - 248.

Gray, J. E. 1852. *Catalogue of Mammalia in the British Museum. Ungulata and Furcipeda* **3**: 277. London.

Grobler, J. H. 1983. Feeding habits of the Cape mountain zebra *Equus zebra zebra* Linn. 1758. *Koedoe* **26**: 159 – 168.

Groves, C. P. 1974. *Horses, Asses and Zebras in the Wild.* Ralph Curtis Books, Hollywood, Florida.

Grzimek, B., Grzimek, B. 1960. A study of game of the Serengeti Plains. *Z. Säugetierk* **25**: 1 – 61.

Guo, S. W., Thompson, E. A. 1992. Performing the exact test of Hardy-Weinberg proportions for multiple alleles. *Biometrics* **48**: 361 – 372.

Hack, M. A., East, R., Rubenstein, D. I. 2002. Status and Action Plan for the Plains Zebra (*Equus burchelli*). In: *Equids: Zebras, Asses and Horses: Status Survey and Conservation Action Plan.* (ed. Moehlman, P. D.) IUCN/SSC, Gland, Switzerland.

Harley, E. H. 2001. DaPSA 4.91, a program for DNA and protein sequence analysis. Wildlife Genetics Unit, University of Cape Town, South Africa.

Harley, E. H. 2002. AGaR_{ST} 2.8, a program for calculating allele frequencies, G_{ST} and R_{ST} from microsatellite data. Wildlife Genetics Unit, University of Cape Town, South Africa.

Hartl, D. L., Clark, A. G. 1989. *Principles of Population Genetics.* Third Edition. Sinauer, Sunderland, MA, USA.

Hartl, G. B., Pucek, Z. 1994. Genetic Depletion in the European Bison (*Bison bonasus*) and the Significance of electrophoretic heterozygosity for conservation. *Cons. Biol.* **8**: 167 - 174.

Hasegawa, M., Kishino, H., Yano, K. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**: 160 - 174.

Hedrick, P. W. 1999. Highly variable loci and their interpretation in evolution and conservation. *Evolution* **53**: 313 – 318.

Henderson, S. T., Petes, T. D. 1992. Instability of simple sequence DNA in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **12**: 2749 - 2757.

Hendey, Q. B. 1974. Faunal dating of the late Cenozoic of southern Africa, with special reference to the Carnivora. *Quatern. Res.* **4**: 149 – 161.

Hennig, W. 1966. *Phylogenetic Systematics.* University of Illinois Press, Urbana, IL.

Higuchi, R. G., Wrischnik, L. A., Oakes, E., George, M., Tong, B., Wilson, A. C. 1987. Mitochondrial DNA of the extinct quagga: relatedness and extant of postmortem change. *J. Mol. Evol.* **25**: 283 - 287.

Hilton-Taylor, C. 2000. *2000 IUCN Red List of Threatened Species.* IUCN, Gland, Switzerland and Cambridge, UK.

Hopman, T. J., Han, E. B., Story, M. R., Schug, M. D., Aquadro, C. F., Bowling, A. T., Murray, J. D., Caetano, A. R., Antczak, D. F. 1999. Equine dinucleotide repeat loci COR001-COR020. *Anim Genet.* **30**: 225 - 226.

Jarne, P., Lagoda, P. J. L. 1996. Microsatellites, from molecules to populations and back. *Trends. Ecol. Evol.* **11**: 424 - 430.

Jeffreys, A. J., Royle, N. J., Wilson, V., Wong, Z. 1988. Spontaneous mutation rates to new length alleles at tandem-repetitive hypervariable loci in human DNA. *Nature* **332**: 278 - 281.

Jones, M. L. 1993. Longevity of ungulates in captivity. *International Zoo Yearbook* **31**: 159 – 169.

Joubert, E. 1971. Ecology, behaviour and population dynamics of the Hartmann's mountain zebra, *Equus zebra hartmannae* Matshie, 1898 in South West Africa. D.Sc. thesis, University of Pretoria, Pretoria, South Africa.

Joubert, E. 1974. Notes on reproduction in the Hartmann's mountain zebra, *Equus zebra hartmannae* in South West Africa. *Madoqua* ser. **8**: 31 – 35.

Karesh, W. B., Smith, F., Frazier-Taylor, H. 1987. A remote method for obtaining skin biopsy samples. *Cons. Biol.* **3**: 261-262.

Keller, L. F., Arcese, P., Smith, J. N. M., Hochachka, W. M., Stearns, S. C. 1994. Selection against inbred song sparrows during a natural population bottleneck. *Nature* **372**: 356 – 357.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**: 111 - 120.

Kimura, M. 1983. *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge, UK.

Kimura, M., Crow, J. F. 1963. The measurement of effective population number. *Evolution* **17**: 279 – 288.

Kimura, M., Crow, J. F. 1964. The number of alleles that can be maintained in a finite population. *Genetics* **49**: 725 - 738.

Kingdon, J. 1997. *The Kingdon field guide to African mammals*. Academic press London.

Kingswood, S. C., Kumamoto, A. T., Charter, S. J., Jones, M. L. 1998. Cryptic chromosomal variation in suni *Neotragus moschatus* (Artiodactyla, Bovidae). *Anim. Cons.* **1**: 95 – 100.

Kumar, S., Tamura, K., Jakobsen, I. B., Nei, M. 2001. MEGA2: Molecular Evolutionary Genetics Analysis software, Arizona State University, Tempe, Arizona, USA.

Lande, R. 1993. Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *Am. Nat.* **142**: 911 – 927.

Lande, R. 1995. Breeding plans for small populations based on the dynamics of quantitative genetic variance. In: *Population Management for Survival and Recovery*. (eds Ballou, J. D., Gilpin, M., Foose, T. J.). pp. 318 – 340. Columbia University Press, New York.

Langella, O. 1999. Populations. A software program to calculate distances and create trees. Centre National de la Recherche Scientifique, France.

Lazary, S., Gerber, H., Glatt, P. A., and Straub, R. 1985. Equine leucocyte antigens in sarcoid-affected horses. *Equine Vet. Journ.* **17**: 283 – 286

Lazary, S., Marti, E., Szalai, G., Gaillard, C., and Gerber, H. 1994. Studies on the frequency and associations of equine leucocyte antigens in sarcoid and summer dermatitis. *Anim. Genet.* **25**: 75 – 80

Lehmann, T., Hawley, W. A., Grebert, H., Collins, F. H. 1998. The effective population size of *Anopheles gambiae* in Kenya: implications for population structure. *Mol. Biol. Evol.* **15**: 264 – 276.

Levinson, G., Gutman, G. A. 1987. High frequency of short frameshifts in poly-CA/GT tandem repeats borne by bacteriophage M13 in *Escherichia coli* K-1 2. *Nucl. Acids Res.* **15**: 5323 - 5338.

Linnaeus, C. 1758. *Systema naturae* 10th edition **1**: 74.

Lloyd, P. H. 1984. The Cape mountain zebra 1984. *Afr. Wildlife* **38**: 144 – 149.

Lloyd, P. H., Rasa, O. A. E. 1989. Status, reproductive success and fitness in Cape mountain zebra (*Equus zebra zebra*). *Behav. Ecol. and Sociobiol.* **25**: 411 – 420.

Luikart, G., Comuet, J. M., Allendorf, F. W., Sherwin, W. B. 1998. Distortion of allele frequency distributions provides a test for recent population bottlenecks. *J. Hered.* **89**: 238 - 247.

Lynch, M., Gabriel, W. 1990. Mutation load and the survival of small populations. *Evolution* **44**: 1725 - 1737.

MacFadden, B. J., Hulbert, R. C. 1988. Explosive speciation at the base of the adaptive radiation of Miocene grazing horses. *Nature* **336**: 466 – 469.

MacFadden, B. J., Solounias, N., Cerling, T. E. 1999. Ancient diets, ecology, and extinction of 5-million-year old horses from Florida. *Science* **283**: 824 - 827.

Mann, M. 1990. *Zambezi: Journey of a River.* Southern Book Publishers, Midrand.

Marklund, S., Ellegren, H., Eriksson, S., Sandberg, K., Andersson, L. 1994. Parentage testing and linkage analysis in the horse using a set of highly polymorphic microsatellites. *Anim. Genet.* **25**: 19 – 23.

Martens, A., DeMoor, A. 1996. Equine Sarcoid .1. clinical types, prevalence, epidemiology, etiology and pathogenesis. *Vlaams Diergeneesk Tijdschr.* **65**: 10 – 17.

Matschie, P. 1898. *Sitzungsberichte der Gesellschaft naturforschender Freunde zu Berlin.* pp 174.

Matthee, C. A., Robinson, T. J. 1997. Mitochondrial DNA phylogeography and comparative cytogenetics of the springhare, *Pedetes capensis* (Mammalia: Rodentia). *J. Mamm. Evol.* **4**: 53 - 73.

Matthee, C. A., Robinson, T. J. 1999. Mitochondrial DNA population structure of roan and sable antelope: implications for the translocation and conservation of the species. *Mol. Ecol.* **8**: 227 – 239

Mentis, M.T. 1970. Estimates of natural biomasses of large herbivores in the Umfolozi Game Reserve area. *Mammalia* **34**: 363 - 393.

Meredith, D., Elser, A. H., Wolf, B., Soma, L. R., Donawick, W. J., and Lazary, S. 1986. Equine leukocyte antigens: relationships with sarcoid tumors and laminitis in two pure breeds. *Immunogenet.* **23**: 221 – 225.

Michalakis, Y., Excoffier, L. 1996. A generic estimation of population subdivision using distances between alleles with special reference to microsatellite loci. *Genetics* **142**: 1061 – 1064.

Millar, J. C. G. 1970a. Census of Cape mountain zebras: Part I. *Afr. Wildlife* **24**: 17 – 25.

Millar, J. C. G. 1970b. Census of Cape mountain zebras: Part II. *Afr. Wildlife* **24**: 105 – 114.

Minch, E., Ruiz-Linares, A., Goldstein, D., Feldman, M., Cavalli-Sforza, L. L. 1997. Microsat 1.5d : a computer program for calculating various statistics on microsatellite allele data. Human Population Genetics Laboratory, Stanford University, CA.

Moehlman, P. D. 2002. ed. *Equids: Zebras, Asses and Horses: Status Survey and Conservation Action Plan.* IUCN/SSC, Gland,Switzerland.

Moxon, R. E., Wills, C. 1999. DNA Microsatellites: Agents of Evolution? *Scient. Amer.* **Jan 1999:** 94-99.

Nasir, L., Reid, S. W. 1999. Bovine papillomaviral gene expression in equine sarcoid tumours. *Virus Res.* **61**: 171 - 175.

Nei, M. 1972. Genetic distance between populations. *Am. Nat.* **106**: 283 - 292.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**: 583 - 590.

Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nei, M., Kumar, S. 2000. *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.

Novellie, P., Lindeque, M., Lindeque, P., Lloyd, P. H., Koen, J. 2002. Status and Action Plan for the Mountain Zebra (*Equus zebra*). In: *Equids: Zebras, Asses and Horses: Status Survey and Conservation Action Plan*. (ed. Moehlman, P. D.) IUCN/SSC, Gland, Switzerland.

Oakenfull, E. A., Lim, H. N., Ryder, O. A. 2000. A survey of equid mitochondrial DNA: Implications for the evolution, genetic diversity and conservation of *Equus*. *Conservation Genetics* **1**: 341 – 355.

O'Brien, S. J., Roelke, M. E., Marker, L., Newman, A., Winkler, C. A., Meltzer, D., Colly, L., Evermann, J. F., Bush, M., Wildt, D.E. 1985. Genetic basis for species vulnerability in the cheetah. *Science* **227**: 1428 - 1434.

Ohta, T., Kimura, M. 1973. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Gen. Res.* **22**: 201 – 204.

O'Ryan C., Harley, E. H., Bruford, M. W., Beaumont, M. A., Wayne, R. K., Cherry, M. I. 1998. Microsatellite analysis of genetic diversity in fragmented South African buffalo populations. *Anim. Conserv.* **1**: 85 - 94.

Oustalet, M. E. 1882. *La Nature*. Paris **10**: 12

Paetkau, D., Waits, L. P., Clarkson, P. L., Craighead, L., Strobeck, C. 1997. An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. *Genetics* **147**:1943 – 1957.

Penzhorn, B. L. 1975. Behaviour and population ecology of the Cape mountain zebra *Equus zebra zebra* Linn., 1758 in the Mountain Zebra National Park. Ph.D. thesis, University of Pretoria, Pretoria, South Africa.

Penzhorn, B. L. 1984a. Observations on mortality of free-ranging Cape mountain zebras *Equus zebra zebra*. *S. Afr. J. Wildl. Res.* **14**: 89 – 90

Penzhorn, B. L. 1984b. A long-term study of social organisation and behaviour of Cape mountain zebras *Equus zebra zebra*. *Z. Tierpsychol.* **64**: 97 – 146.

Penzhorn, B. L. 1985. Reproductive characteristics of a free-ranging population of Cape mountain zebra (*Equus zebra zebra*). *J. Reprod. Fert.* **73**: 51 – 55.

Penzhorn, B. L. 1988. *Equus zebra*. *Mammalian Species*. **314**: 1 - 7

Penzhorn, B. L., Lloyd, P. H. 1987. Comparisons of reproductive parameters of two Cape mountain zebra populations. *J. Reprod. Fert. Suppl.* **35**: 661 – 663.

Pringle, J. A. 1982. *The conservationists and killers*. T. V. Bulpin and Books of Africa, Cape Town.

Rau, R. E. 1978. Additions to the revised list of preserved material of the extinct Cape Colony quagga and notes on the relationship and distribution of southern plains zebras. *Ann. S. Afr. Mus.* **77**: 27 – 45.

Rau, R. E. 2002. Are the two mountain zebra subspecies distinguishable? *Mainly about Animals* **54**: 20 – 21.

Raymond, M., Rousset, F. 1995a. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Heredity* **86**:248 - 249

Raymond, M., Rousset, F. 1995b. An exact test for population differentiation. *Evolution* **49** :1280 - 1283.

Reid, S. W., Gettinby, G., Fowler, J. N., Ikin, P. 1994. Epidemiological observations on sarcoid in a population of donkeys (*Equus asinus*). *Vet. Rec.* **134**: 207 - 211

Rice, W. R. 1989. Analysing tables of statistical tests. *Evolution* **43**: 223 – 225.

Roelke, M. E., Martenson, J. S., O'Brien, S. J. 1993. The consequences of demographic reduction and genetic depletion in the endangered Florida panther. *Curr. Biol.* **3**: 340 - 350.

Rogers, A. R., Harpending, H. 1992. Population growth makes waves in the distribution of pair-wise genetic differences. *Mol. Biol. Evol.* **9**: 552 - 569.

Rooney, A. P., Honeycutt, R. L., Davis, S. K., Derr, J. N. 1999. Evaluating a putative bottleneck in a population of bowhead whales from patterns of microsatellite diversity and genetic disequilibria. *J. Mol .Evol.* **49**: 682 – 690.

Rozas, J., Rozas, R. 1999. DNAsp version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* **15**: 174 - 175.

Saitou, N., Nei, M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**: 406 - 425.

Saitou, N., Imanishi, M. 1989. Relative efficiencies of the Fitch-Margolish, maximum parsimony, maximum-likelihood, minimum-evolution and neighbour-joining methods of phylogenetic tree reconstruction in obtaining the correct tree. *Mol. Biol. Evol.* **6**: 514 – 525.

Sambrook, J, Fritsch, E. F., Maniatis, T. (eds). 1989. *Molecular Cloning - A Laboratory Manual*. 2nd edition. Cold Spring Harbour Laboratory Press.

Schneider, S., Roessli, D., Excoffier, L. 2000. Arlequin ver. 2.000: Software for population genetic data analysis. Genetics and Biometry Laboratory. University of Geneva, Switzerland.

Schloetterer, C., Tautz, D. 1992. Slippage synthesis of simple sequence DNA. *Nucl. Acids Res.* **20**: 211 - 215.

Schmidt, H. A., Strimmer, K., Vingron, M., von Haeseler, A. 2000. Tree-puzzle, version 5.0, maximum likelihood analysis for nucleotide, amino acid, and two-state data. Theoretical Bioinformatics, DKFZ, Heidelberg, Germany.

Skead, C. J. 1956. The zebra of the Kammanassie Mountains. *Cape Department of Nature Conservation, Annual Report* **13**: 19 – 22.

Skinner, J. D., Smithers, R. H. N. 1990. *The mammals of the southern African subregion*. University of Pretoria.

Slatkin, M. 1985. Rare alleles as indicators of gene flow. *Evolution* **39**: 53 - 65.

Slatkin, M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* **47**: 264 - 279.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequency. *Genetics* **139**: 457 – 462.

Smuts, G.L. 1972. Seasonal movements, migration and age determination of Burchell's zebra (*Equus burchelli antiquorum*, H. Smith, 1841) in the Kruger National Park (30/11/72 to 4/12/72). M.Sc. dissertation, University of Pretoria, Pretoria, South Africa.

Smuts, G.L. 1974. Growth, reproduction and population characteristics of Burchell's zebra in the Kruger National Park. Ph.D. thesis, University of Pretoria, Pretoria, South Africa.

Sokal, R. R., Sneath, P. H. A. 1963. *Principles of Numerical Taxonomy*. W. H. Freeman and Co., San Francisco.

Spencer, C.C., Neigel, J.E., Leberg, P.L. 2000. Experimental evaluation of the usefulness of microsatellite DNA for detecting demographic bottlenecks. *Mol. Ecol.* **9**: 1517 – 1528.

Strimmer, K., A. von Haeseler. 1996. Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* **13**: 964 - 969.

Studer, U., Marti, E., Stornetta, D., Lazary, S., Gerber, H. 1997 Treatment of equine sarcoid with an unspecific immunostimulant - remarks on the epidemiology and the spontaneous regression of sarcoids. *Schweizer Archiv fur Tierheilkunde* **139**: 385 – 391

Suzuki, D. T., Griffiths, A. J. F., Miller, J. H., Lewontin, R. C. 1989. In *An Introduction to Genetic Analysis 4th ed.* W.H. Freeman p.704.

Swinburne, J. E., Marti, E., Breen, M., Binns, M. M. 1997. Characterisation of twelve new horse microsatellite loci: AHT 12 – AHT 23. *Anim. Genet.* **28**: 453.

Swofford, D. L. 1998. PAUP*. Phylogenetic Analysis Using Parsimony and Other Methods. Version 4. Sinauer, Sunderland, MA, USA.

Taberlet, P., Camarra, J. J., Griffin, S., Uhres, E., Hanotte, O., Waits, L. P., Dobois-Paganon, C., Burke, T., Bouvet, J. 1997. Noninvasive genetic tracking of the endangered Pyrenean brown bear population. *Mol. Ecol.* **6**: 869 – 876

Takezaki, N., Nei, M. 1996. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* **144**: 389 – 399.

Tautz, D., Renz, M. 1984. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucl. Acids Res.* **12**: 4127 - 4138.

Tozaki, T., Kakoi, H., Mashima, S., Hirota, K., Hasegawa, T., Ishida, N., Miura, N., Tomita, M. 2000. The isolation and characterization of 18 equine microsatellite loci, TKY272-TKY289. *Anim Genet.* **31**:149 - 150.

Troy, C. S., MacHugh, D. E., Bailey, J. F., Magee, D. A., Loftus, R. T., Cunningham, P., Chamberlain, A. T., Sykesk, B. C., Bradley, D. G. 2001. Genetic evidence for Near-Eastern origins of European cattle *Nature* **410**: 1088 – 1091.

Turner, M. I. M., Watson, R. M. 1965. Game management and research by aeroplane. *Oryx* **8**: 13 – 22.

Valsecchi, E., Palsbøll, P., Hale, P., Glockner-Ferrari, D., Ferrari, M., Clapham, P., Larsen, F., Matilla, D., Sears, R., Sigurjonsson, J., Brown, M., Corkeron, P., Amos, B. 1997. Microsatellite genetic distances between oceanic populations of the humpback whale (*Megaptera novaeangliae*). *Mol. Biol. Evol.* **14**: 355 – 362.

van der Merwe, N. J. 1962. The position of nature conservation in South Africa. *Koedoe* **5**: 1 - 122

van Haeringen, W. A., van de Goor, L. H. P., van der Hout, N., Lenstra, A. 1998. Characterization of 24 microsatellite loci. *Anim. Genet.* **29**: 153 – 156.

Vila, C., Leonard, J. A., Götherström, A., Marklund, S., Sandberg, K., Lidén, K., Wayne, R. K., Ellegren, H. 2001. Widespread origins of domestic horse lineages. *Science* **291**: 474 – 477.

Waser, N. M. 1993. Population structure, optimal outbreeding, and assortative mating in angiosperms. In *The natural history of inbreeding and outbreeding* (ed. Thornhill, N. W.) University of Chicago Press, Chicago

Waser, P. M., Strobeck, C.1998. Genetic signatures of interpopulation dispersal. *Trends in Ecology and Evolution* **13**: 43 - 44.

Weir, B. S. 1996. *Genetic Data Analysis II: Methods for discrete population genetic data.* Sinauer, Sunderland, MA, USA.

Weir, B. S., Cockerham. C. C., 1984. Estimating F-statistics for the analysis of population structure. *Evolution* **38**: 1358 - 1370.

Whitehouse, A.M., Harley, E.H. 2001. Post-bottleneck genetic diversity of elephant populations in South Africa, revealed by microsatellite analysis. *Mol. Ecol.* **10**: 2139-2149.

Wilson, G. A., Strobeck, C. 1999. Genetic variation within and relatedness among wood and plains bison populations. *Genome* **42**: 483-96

Wiseman, R., O'Ryan, C., Harley, E.H. 2000. Microsatellite analysis reveals that domestic cat (*Felis catus*) and southern African wild cat (*F. lybica*) are genetically distinct. *Anim. Cons.* **3**: 221-228.

Woods, D. H. 1960. Mountain zebras. *J. Mount. Club S. Afr.* **63**: 4 – 9

Woods, D. H. 1962. On zebras in the Couga Mountains *J. Mount. Club S. Afr.* **65**: 16 – 17

Wright, S. 1931. Evolution in Mendelian populations. *Genetics* **16**: 97-159.

Wright, S. 1943. Isolation by distance. *Genetics* **28**: 139 –156.

Wright, S. 1969. *Evolution and the genetics of populations. Vol. 2. The theory of gene frequencies.* University of Chicago Press, Chicago.